

Model Based Defect Reconstruction in Ultrasonic Non-Destructive Testing

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen
Doktorgrades
“Doctor rerum naturalium”
der Georg-August-Universität Göttingen

vorgelegt von
Florian Boßmann
aus Emmerich

Göttingen 2013

Betreuungsausschuss

- Prof. Dr. Gerlind Plonka-Hoch, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen
- Prof. Dr. Thorsten Hohage, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission

- Referentin:
Prof. Dr. Gerlind Plonka-Hoch, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen
- Korreferent:
Prof. Dr. Russell Luke, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission

- Prof. Dr. Thorsten Hohage, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen
- Prof. Dr. Anja Sturm, Institut für Mathematische Stochastik, Georg-August-Universität Göttingen
- Prof. Dr. Stephan Waack, Institut für Informatik, Georg-August-Universität Göttingen
- Prof. Dr. Ingo Witt, Mathematisches Institut, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 25.10.2013

“ He that breaks a thing
to find out what it is
has left the path of wisdom.”

J. R. R. TOLKIEN

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Prof. Dr. Gerlind Plonka-Hoch for her exhaustless support during the preparation of this thesis. She had not only always time for helpful discussions on my work but also contributes several useful hints and suggestions. Moreover, starting with the very first terms of my diploma study, she introduced me into the interesting field of signal- and image-processing. Furthermore, I express my gratitude to Prof. Dr. Thorsten Hohage, who accompanies me as co-advisor during all the time, and to Prof. Dr. Russell Luke, who was so kind to undertake the task of being the co-referee of my thesis.

Besides, I would like to thank the complete group “Mathematical Signal- and Image-Processing” for numerous inspiring conversations in a good working atmosphere, not only on the topic of this work. Especially, I would like to mention Jun.-Prof. Dr. Felix Kraemer, for establishing the tea break in our group that enjoys great popularity, and David James, Stefan Loock, and Katrin Wannewetsch for proof-reading parts of this work.

This work would not have been possible without the support of several organisations. The subject of this thesis and the data used within it was provided by Mannesmann Salzgitter Forschung GmbH. Specifically, I would like to thank the members of the group of non-destructive testing Alfred Graff, Dr. Oliver Nemitz, Dr. Thomas Orth, and Dr. Till Schmitte for their continuous support with data and information. I appreciate BMWI for the financial support of this cooperative project MoMuz. Furthermore, I am grateful that I had the opportunity to be member of the SFB755 which not only supplied financial support but also organised workshops in an inspiring atmosphere. I would also like to mention BMBF for the opportunity to continue my research in the recently established joint research project ZeMat.

Last but not least I want to thank my family and friends. They provided distraction when it was needed and may forgive me for all the times I neglected them. I am thankful to Michael and Arndt who, although outside the field of natural science, had the heart to proof-read parts of my thesis. I am deeply grateful to have Stephanie at my side always giving me support and appreciation.

Contents

1	Introduction	1
2	Ultrasonic Non-Destructive Testing (NDT)	3
2.1	Principles of Ultrasound	3
2.1.1	Oscillations and Waves	3
2.1.2	Huygens Principle	5
2.1.3	Reflection, Refraction, Diffraction	7
2.2	Ultrasonic NDT Methods	8
2.2.1	The Piezoelectric Effect	9
2.2.2	Measuring with Ultrasound	9
2.2.3	Time of Flight Diffraction (ToFD)	12
2.2.4	Wall Thickness Measurement	13
3	Defect Reconstruction: State of the Art	15
3.1	Reconstruction Methods in NDT	15
3.1.1	Applying Gates	15
3.1.2	Total Focusing Method (TFM)	16
3.1.3	Further Reconstruction Methods	18
3.2	Inverse Scattering	18
3.2.1	The Direct and Inverse Problem	19
3.2.2	Reconstruction Methods for Complete Data	22
3.2.3	The Limited Aperture Case	24
3.3	Other Fields of Research	26
4	Model	29
4.1	The Continuous Model	29
4.1.1	The Geometrical Approach	30
4.1.2	Wave Convolution	40
4.2	Discretization	42
4.3	Model Validation	43
4.3.1	Comparison to an Inverse Scattering Approach	44
4.3.2	Numerical Validation	46
5	Inversion - an Overview	53

6	Inversion 1 - Deconvolution	59
6.1	Sparse Deconvolution	60
6.1.1	(Orthogonal) Matching Pursuit	60
6.1.2	MP and OMP for Sparse Deconvolution of NDT Data	63
6.1.3	Some Comments on Approximation Properties of MP	67
6.2	Blind Deconvolution	69
6.2.1	Calculation of the Starting Parameters	70
6.2.2	Parameter Update by Newton-Iteration	73
6.3	Summary	74
7	Inversion 2 - Clustering	77
7.1	Cluster Analysis	78
7.2	Clustering Algorithms for Separated Defect Peaks	79
7.3	Clustering Algorithms for Intersecting Defect Peaks	83
7.4	Summary	86
8	Inversion 3 - Defect Reconstruction	89
8.1	Reconstruction of the Center Point	92
8.1.1	A Heuristic Method for ToFD	94
8.1.2	A Heuristic Method for Wall Thickness Measurement	96
8.1.3	A General Heuristic Approach	97
8.2	Spiderweb Algorithm for Defect Reconstruction	99
8.2.1	Reconstruction as Minimization Problem	99
8.2.2	Supervised Minimization	108
8.2.3	Unsupervised Minimization	117
8.3	Summary	124
9	Numerical Results	127
9.1	Approximation Results for OMP and Newton-HOMP	127
9.2	Clustering Results for DBSCAN and DBCLAN	132
9.3	Reconstruction with the Spiderweb Algorithm	138
9.4	Numerical Results for the Complete Inversion Method	141
9.4.1	Simulated ToFD Pore	141
9.4.2	Real ToFD Data	150
9.4.3	Simulated Back Wall Defect	152
9.4.4	Real Wall Thickness Measurement Data	158
10	Conclusions	161
	Bibliography	163
	Curriculum Vitae	171

1 Introduction

Although the history of iron converting can be traced back long before Christ [1], steel had over the centuries only a minor role as material. It was in most cases only used in war techniques. This changed around the 19th century with the beginning of the industrial revolution, inventions like the steam engine (1712 by Thomas Newcomen) and the railway, the market for steel grew rapidly. In the middle of the 19th century faster production methods were developed by Henry Bessemer (1855) and Sidney Gilchrist Thomas (1855) to cover the demand. But with this new market also new requirements were made on the produced components. They had to stand high pressure or enormous forces. As a consequence, there were too many accidents like boiler explosions, derailling or even bridge collapses, often caused by only little defects in some components. Reliable methods for material testing had to be developed.

Until the end of the 19th century there were no effective methods for testing steel components. Although one was able to evaluate several material properties, it was hardly possible to find defects inside the material. This changed with the discovery of x-rays by Wilhelm Conrad Röntgen in 1895. But while x-rays came to immediate use in medical science the use in non-destructive material testing was problematic. The equipment was big and expensive, the testing process was time-consuming and dangerous, and only thin material could be tested. Due to these facts the x-ray non-destructive testing did not become important until 1930. With the discovery of the piezoelectric effect (page 9) by Pierre Curie in 1883 the foundation for ultrasonic testing was laid. Ultrasonic testing did not possess the disadvantages of x-rays but it still took 60 years until the first feasible equipment was produced.

The historical introduction above shown is based on [2]. Nowadays there are many different non-destructive testing methods known based on x-rays, ultrasound, eddy current or other techniques. However, while non-destructive testing techniques are still enhanced, the automation of the complete steel production process presents new problems. The data given by the non-destructive testing method has to be analysed in a fast and reliable way without slowing down the production.

In this thesis a fast and stable algorithm for ultrasonic data is introduced that is able to reconstruct main features of a defect e.g. position, form, size and type. For this purpose, Chapter 2 will shortly introduce the technique of ultrasonic non-destructive testing. Chapter 3 will touch upon the currently known methods from the field of non-destructive testing as well as the ideas from close-by fields like inverse scattering. The remaining chapters of this thesis will describe our new approach to the problem. Chapter 4 will

provide a simple model for ultrasonic non-destructive testing and Chapters 5 to 8 will focus on a new inversion method for this model. Chapter 9 will present numerical results as well as a comparison to the methods shown in Chapter 3. Finally, the last chapter will conclude this thesis.

2 Ultrasonic Non-Destructive Testing (NDT)

In this chapter we give a brief introduction to the physical and technical background of the methods considered in this work. The first section covers the emergence and behaviour of ultrasound in steel. For a better understanding this is done in an illustrative way following [3]. For more information we refer to [4–6] and references therein or to [7] for a detailed mathematical introduction. We would also like to mention [8] where a short introduction into ultrasound and its applications in non-destructive testing is given including a number of tables listing material constants needed in this work.

The second section of this chapter will illuminate the generation and measuring of ultrasound. The different methods of ultrasonic testing will be introduced and their mode of operation will be illustrated.

2.1 Principles of Ultrasound

2.1.1 Oscillations and Waves

Before we can understand the propagation of ultrasound in materials it is necessary to understand oscillations and waves. To demonstrate oscillation we consider an object that is elastically fixed at a point. If we now stretch the object away from its origin and release it, it will be pulled back to the origin because of the elastic fix. Furthermore, due to the speed it has at the origin, it will move across it and drift away stretching the elastic fix again (Figure 2.1). Note that we do not consider a loss of energy in this chapter. Oscillations and waves usually decrease over time by effects like reflection, refraction, friction and others but this decrease will be specified later on.

In solid materials like steel, the particles are not fixed to a certain point but elastically bounded by their neighbouring particles (Figure 2.2). If one particle starts to oscillate, it will affect its neighbours which will now also start to oscillate. Thus a wave propagates through the particle grid. Note that a time shift appears between the oscillation of two neighbouring particles caused by the elastic bound between them. In order to illustrate this effect, we will consider two special cases. In the first case we look at the behaviour of the grid in perpendicular direction to the oscillation. Therefore we consider the first column to oscillate in vertical direction while all vertical connections stay fixed

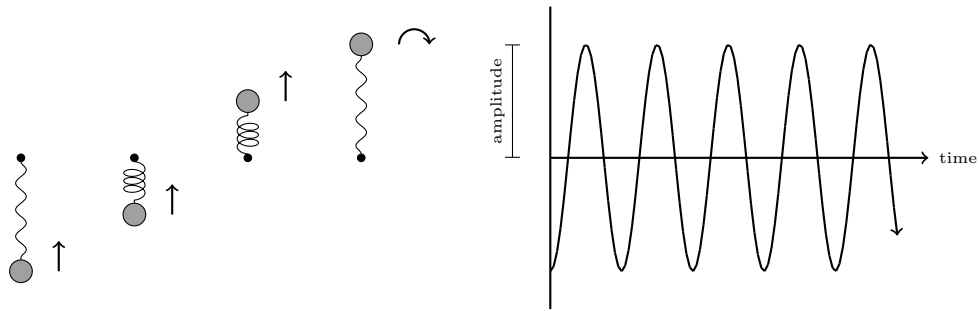


Fig. 2.1: Oscillation of a elastically fixed object (left) and expansion plotted over time (right).

(Figure 2.3). This effect is the so-called transverse wave. In the second case, we consider an oscillation in horizontal direction while all vertical connections still stay fixed (Figure 2.4). This wave is called the longitudinal wave. In practice the appearing waves will always be a combination of both types. Note that the propagation shown in Figures 2.3 and 2.4 only holds far away from boundaries.

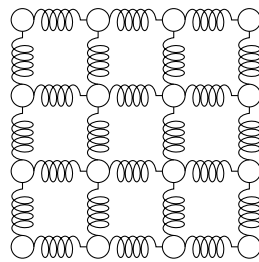


Fig. 2.2: Sketch of a particle grid.

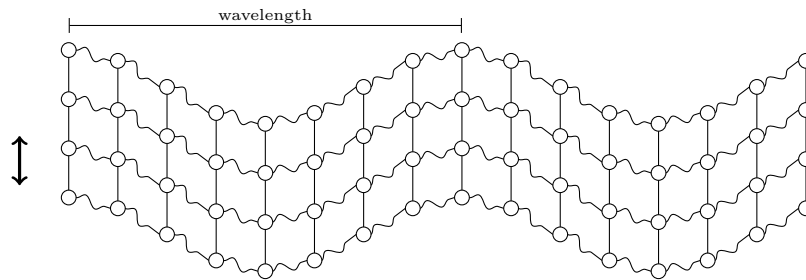


Fig. 2.3: Wave propagation perpendicular to the oscillation.

Let us now introduce some ultrasound parameters needed for later on discussions. For this purpose, we only consider isotropic homogeneous materials in this work. For non-homogeneous materials we refer to [9]. In isotropic homogeneous materials the propagation speed of ultrasound is constant and depends on the wave type. In steel, the longitudinal wave has approximately a speed of $5.94 \frac{\text{mm}}{\mu\text{s}}$ while the transversal wave only propagates with $3.29 \frac{\text{mm}}{\mu\text{s}}$ [8]. These values can differ depending on the type of steel,

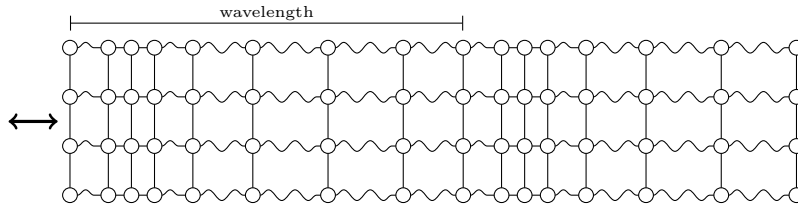


Fig. 2.4: Wave propagation in oscillation direction.

other values might also be found in different literatures. For the applications considered in this work this large speed difference leads to the situation that the transversal waves will usually be measured a long time after the last relevant longitudinal wave has been recorded. Hence we only consider longitudinal waves in our model as done in [10] and denote

$$c = 5.94 \frac{\text{mm}}{\mu\text{s}}. \quad (2.1)$$

The frequency ω of a wave is defined by the oscillations of a particle per second. If the frequency is above 20kHz ($\frac{10^3}{\text{s}}$), it is called an ultrasound wave [5]. For ultrasonic testing ω is normally taken between 0.5Mhz and 10Mhz [3]. Furthermore, we define the amplitude of a wave by the maximum deflection (Figure 2.1) and the wavelength λ by the distance in mm between to points in the wave with the same deflection (Figure 2.3 and 2.4). Note that

$$\omega\lambda = c \quad (2.2)$$

holds [3]. Now we have analysed the emergence and behaviour of ultrasound far away the boundary. To understand the phenomena of reflection, refraction and diffraction, we first have to introduce the Huygens principle in the next subsection.

2.1.2 Huygens Principle

In 1690 Christiaan Huygens published his work "Traité de la lumière" (Treatise on light [11]) where he showed the effect of reflection and refraction with a simple principle that was named after him later. Although his work was mainly applied to light waves, the principle can be used for all types of waves and also explains the effect of diffraction.

”There is the further consideration in the emanation of these waves, that each particle of matter in which a wave spreads, ought not to communicate its motion only to the next particle which is in the straight line drawn from the luminous point, but that it also imparts some of it necessarily to all the others which touch it and which oppose themselves to its movement. So it arises that around each particle there is made a wave of which that particle is the centre. ...” [11]

For a better understanding of this principle we first denote a wave front [5, 12] as those neighbouring points with same phase, this means with same displacement to their origin. In Figure 2.4, all particles in one column build a wave front while the distance between two wave fronts of the same phase is the wavelength. Far away from boundaries all particles that start to oscillate at the same time build a wave front. For isotropic homogeneous media this means all particles with the same distance to the source build a wave front. Figure 2.5 shows typical wave fronts for different sources. Note that effects at boundaries like reflection can cause interference of ultrasound waves and thereby influence the shape of the wave fronts.

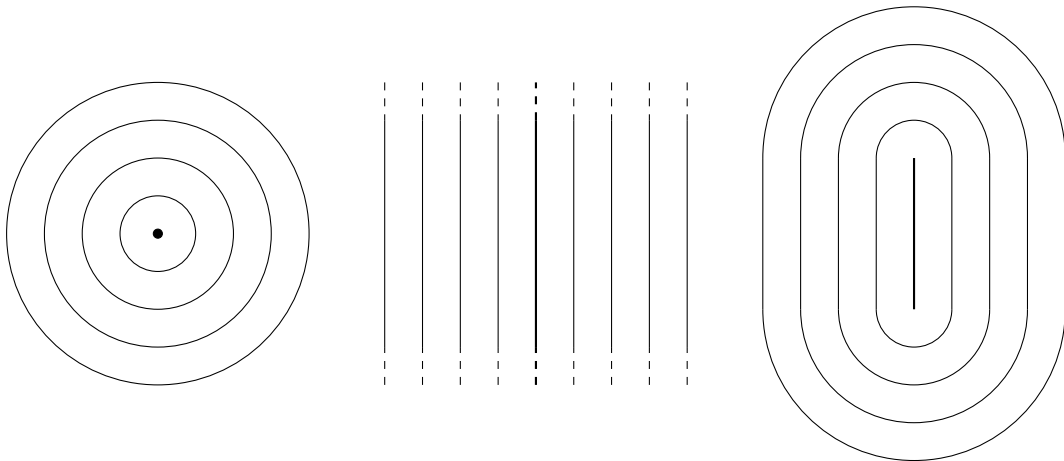


Fig. 2.5: Typical wave fronts (thin) for different sources (thick).

We now understand “each particle of matter in which a wave spreads” as the wave front with the largest distance to the source. This means the boundary between oscillating and not yet oscillating particles. Huygens principle can be illustrated as shown in Figure 2.6, where each particle of the old wave front is the center of a so-called elementary wave, a half sphere in propagation direction with radius depending on the ultrasound speed and the considered time step. The closure of all elementary waves is the new wave front.

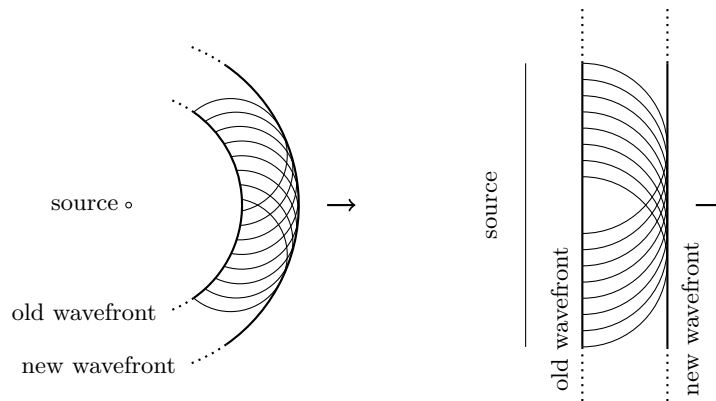


Fig. 2.6: Huygens principle for a point source (left) and a plane wave (right).

2.1.3 Reflection, Refraction, Diffraction

With the help of Huygens principle we will now demonstrate the effect of reflection, refraction and diffraction at boundaries, see also [3, 5, 13, 14]. In the last subsection the principle was shown for waves far away from the boundary. If a wave approaches a boundary we can observe different effects depending on the type of the boundary.

The first effect appears when a wave hits a (straight) surface. The elementary waves at the boundary build a new wave front that propagates as a reflected wave. Note that at boundaries also mode conversion may occur. This means, a part of the incoming wave is converted to a different wave type, i.e. an incoming wave will cause a reflected transversal and longitudinal wave. We only consider longitudinal waves in our model so we do not consider the effect of mode conversion. Figure 2.7 (top) shows the reflection of an incoming ultrasound wave without mode conversion. Because the ultrasound speed stays the same, the angle of reflection is equal to the angle of incidence. The mode converted elementary waves have a different speed and the wave is reflected at a different angle similarly to the effect we will see at refraction.

Next we will analyse diffraction. This effect takes place at the edges of boundaries. Here it seems as if ultrasound can travel round the defect. At first glance this seems to be unlikely but remember that the new wave front is the closure of all elementary waves. As soon as a wave passes the edge, these elementary waves are no longer restricted by the boundary and the wave propagates around the edge. Figure 2.7 (middle) shows this effect for a wave propagating perpendicular to the boundary.

While reflection and diffraction are effects that take place in the same material, refraction explains the transfer of ultrasound between two bordering materials. An ultrasound wave that propagates to a boundary will also cause elementary waves in the bordering material. These elementary waves build a new wave front. Because the ultrasound speed

in both materials may differ, the refracted wave propagates in a different direction. The connection between incoming and refracted angle can be explained by Snell's law [3, 5]

$$\frac{\sin \alpha}{\sin \beta} = \frac{c_1}{c_2},$$

where c_1 and c_2 is the ultrasound speed in the materials and α and β are the propagating angles (see Figure 2.7 (bottom)). Snell's law can also be applied to the reflection angle of a mode converted wave. Note that for $(c_2/c_1) \sin \alpha > 1$, there is no possible solution for β . This means no refraction (or mode converted reflection) takes place. The angle α with $(c_2/c_1) \sin \alpha = 1$ is called critical angle [5]. Figure 2.7 (bottom) shows the effect of refraction.

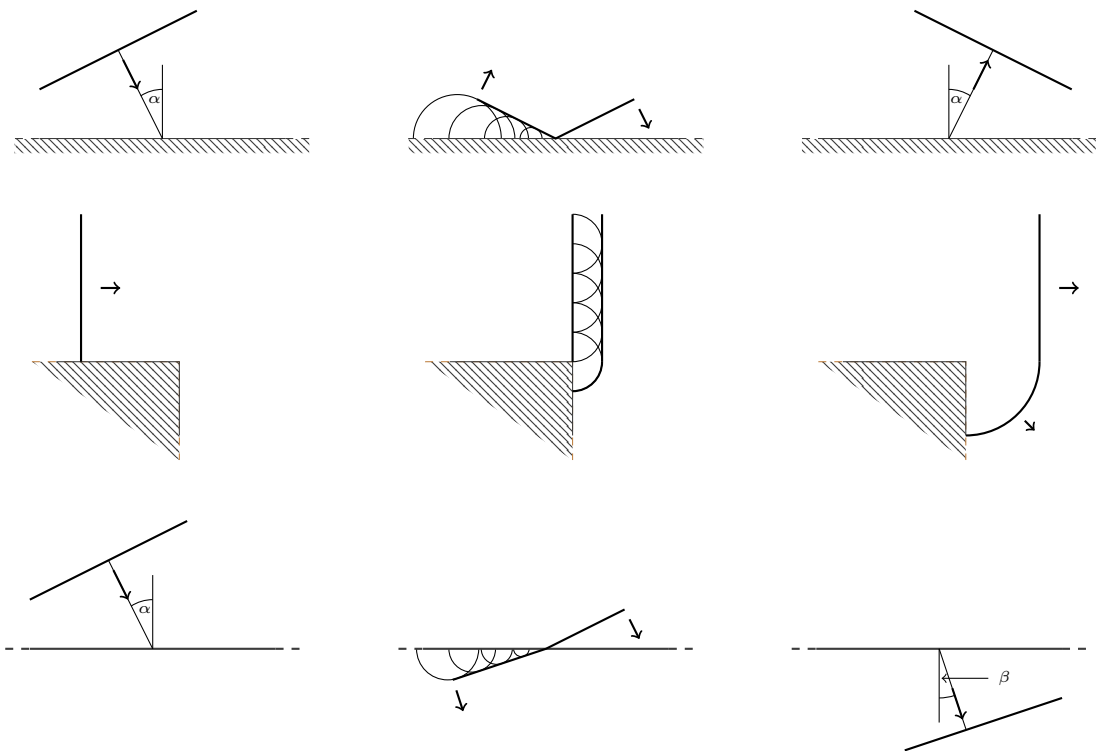


Fig. 2.7: Huygens principle for reflection (top), diffraction (middle) and refraction (bottom).

2.2 Ultrasonic NDT Methods

In this section we will discuss the use of ultrasound for non-destructive testing methods. Therefore we first have to learn how to excite and measure ultrasonic waves. This can be done with the piezoelectric effect explained in the next subsection. Later in this section

we will introduce different testing techniques and see how they take advantage of the different effects reflection, refraction and diffraction.

2.2.1 The Piezoelectric Effect

The piezoelectric effect is very important in the exciting and measuring of ultrasound. This section will only deliver a short insight to this phenomenon. For detailed information we refer to [3,4] and the references therein. A short review can also be found in [5].

To test a component part with ultrasound, one needs to be able to create an ultrasonic impulse with desired properties like direction, amplitude and frequency. The most common way to do so, is to use piezo elements like barium titanate [3]. These elements have the property to change in size if a voltage is applied. Depending on the voltage direction it increases or decreases. This is called the reverse piezoelectric effect. Therewith it is possible to regulate the amplitude and frequency of the generated ultrasound directly.

While the reverse effect is used to create ultrasound, the direct piezoelectric effect can be used to measure it. If mechanical stress is applied to a piezo element, it creates a voltage that can be measured. That means, if an ultrasonic wave hits the piezo element the amplitude and frequency can be measured by measuring the created voltage. Figure 2.8 illustrates the piezoelectric effect, whereat both directions are possible: Mechanical stress is applied and voltage is created or voltage is applied and the piezo element changes in size.

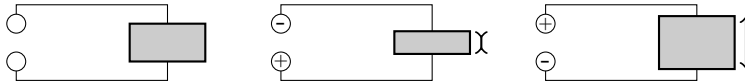


Fig. 2.8: Piezoelectric effect.

Most of the probes used nowadays are mainly based on the piezoelectric effect. A piezo element is attached to an electric circuit. Thus the probe can be used as an emitter or receiver by using the reverse or direct piezoelectric effect. We would like to mention at this point that the probes used nowadays also have some further techniques to focus the generated ultrasound, see [3,4].

2.2.2 Measuring with Ultrasound

This subsection will give a short introduction into the general set-up of ultrasonic non-destructive testing. For detailed information see [3,4,15]. Understanding the functionality of ultrasound probes, this general set-up can be explained as shown in Figure 2.9. An electric impulse is applied to the emitter which converts it into an ultrasonic impulse. This impulse is inducted into the component. Depending on the inner structure of the component, parts of the ultrasound itself and reflections and diffractions of it reach the

receiver that converts the signal back into electricity producing the measured data. Normally, a coupling liquid is placed between the probes and the component, but we do not attend to it in this short introduction.

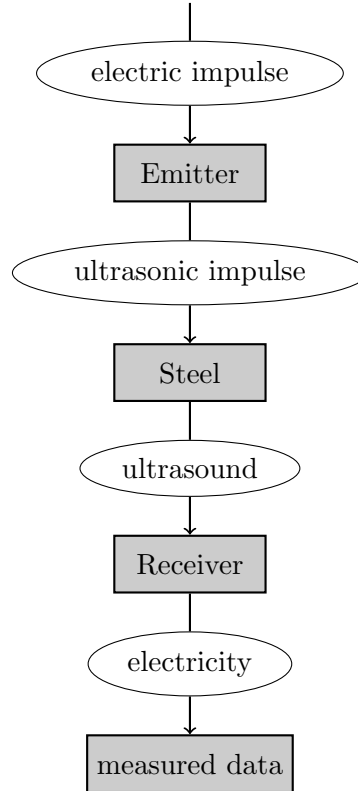


Fig. 2.9: Set-up for ultrasonic testing.

Note that Figure 2.9 assumes only one emitter and receiver which is typical for ultrasonic non-destructive testing. That way, the methods are more flexible and transportable. In order to gather a high amount of data, several measurements are made with different probe positions. Thus one speaks about a-scan and b-scan data. The a-scan is the data of one measurement with fixed probes positions, it is a graph of the measured electricity over time (see Figure 2.10 (left)). The b-scan is a two-dimensional matrix where each column contains an a-scan, i.e. it is a two dimensional structure where one dimension is the time and the other dimension describes the probe position. For example, if we move the probes along a line, we get a two-dimensional b-scan as shown in Figure 2.10 (right), where each column stands for a probe position on this line. One fixed column is just the grey-scaled a-scan belonging to this probe position. If we move the probes in more than one direction, we obtain higher dimensional b-scans.

Now we want to classify ultrasonic testing methods into two classes depending on the incoming ultrasound. The first class is called *impulse-echo technique*. Thereby, the emitter and receiver are placed in a way on the component such that reflections or diffractions of

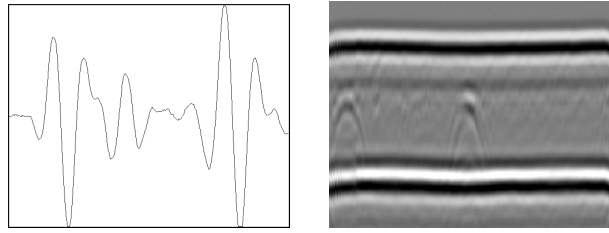


Fig. 2.10: A-scan (left) and b-scan (right).

defects in the component reach the receiver. Figure 2.11 (left) shows a typical example where one probe, emitter and receiver in one, is placed on the surface of the component. The emitted ultrasound will be reflected by the defect and on this way propagates back to the probe. As we will see later, there are also different examples, where emitter and receiver are placed differently to measure diagonal reflections / diffractions.

The second class is the so-called *transmission technique*. Here the receiver is placed on the opposite border of the material such that a possible defect is situated between both probes. Unless there is a defect in the component, the transmitted ultrasound reaches the receiver in an extenuated form (see Figure 2.11 (right)). The impulse-echo technique is much more common because the measured signals contain information about the defect boundary, and hence more conclusions about the defect are possible. For example, the current time of the ultrasound is directly related to the distance between defect and probes. The algorithm presented in this work also needs this information, thus we will now concentrate on impulse-echo techniques.

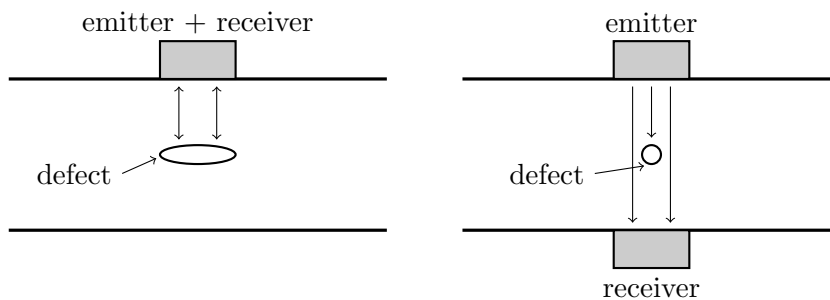


Fig. 2.11: Impulse-echo (left) and transmission technique (right).

One can find plenty of different impulse-echo techniques in the field of non-destructive testing, some of them highly adapted to their testing problem. In this work we will

specifically concentrate on two methods that will be introduced in the next subsections: the Time-of-Flight-Diffraction technique [3, 10, 16–18] and the wall thickness measurement [3, 4].

2.2.3 Time of Flight Diffraction (ToFD)

The first method we want to present is the Time of Flight Diffraction technique. As implied by its name, this technique makes use of the diffraction effect. A typical application is the inspection of welds. It operates with two probes, one emitter and one receiver, that are placed symmetrically to the weld. The emitter transmits an ultrasound impulse that propagates to the weld. If there is a defect in the weld, the diffraction caused by it will propagate to the receiver and hence will be measured. Thus the ToFD method can detect the defect.

Besides the diffraction, there are three other signals that can appear in the measured data: the lateral wave, defect reflections and the back wall echo. The lateral wave is the first signal that arrives at the receiver. It is the ultrasound wave that directly travels from the emitter to the receiver on the shortest way. The back wall echo is the last signal that appears, it is the reflection of the lower boundary of the component. Both, the lateral wave and the back wall echo, appear in almost every measurement and are typically very strong signals. All signals in the meantime are caused by diffractions from defects and material inhomogeneities. In some cases one may also measure reflections from the defect surface, but this signal mostly has a very low amplitude and often disappears in the noise. Figure 2.12 illustrates the ToFD technique.

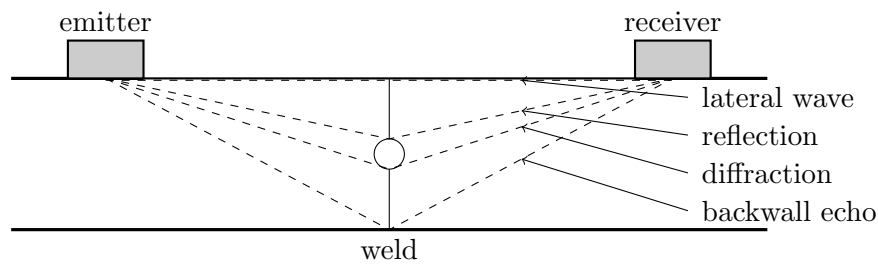


Fig. 2.12: Time of Flight Diffraction method (ToFD).

To inspect the whole weld one needs more than one measurement. Thus after one measurement is done both probes are moved parallel to the weld by a fixed small distance and another measurement is done. That way one gets a two-dimensional b-scan as shown in Figure 2.13. Here every column is an a-scan of a fixed probe position. One can see the typical parabola-like patterns of a defect measurement, the larger the distance between probes and defect the later the diffracted signal arrives.

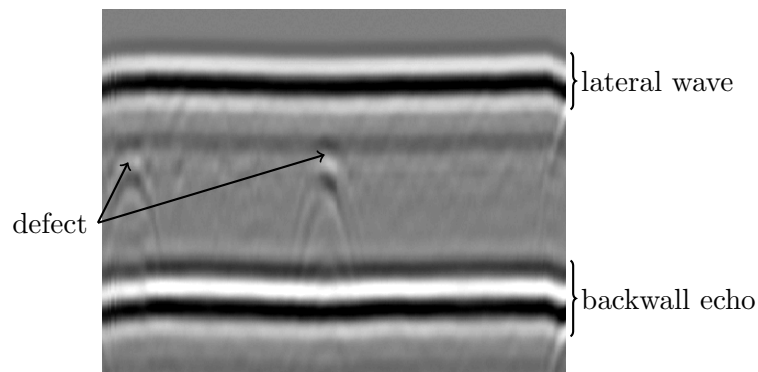


Fig. 2.13: ToFD b-scan.

2.2.4 Wall Thickness Measurement

The second method we would like to introduce is a typical application of an impulse-echo technique with one probe acting as emitter and receiver at once. It seems clear to use non-destructive testing methods to find defects inside a component. But even defects that are on the surface can be hard to find if the surface is not easily accessible. For example, the inner surface of a tube may be hard to analyse. Thus one needs a method that can “see” through the outer surface to inspect the inner one. This can be done with a wall thickness measurement. Therefore a probe, emitter and receiver at once, is placed on the outer surface. It emits ultrasound waves in the direction of the inner surface which reflects them back to the probe. Because there is a direct correlation between the arrival time of the reflection and the wall thickness, one can reconstruct the inner surface. Figure 2.14 shows the set-up of a wall thickness measurement.

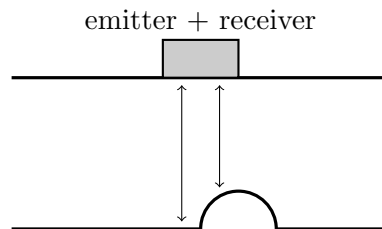


Fig. 2.14: Wall thickness measurement.

In order to reconstruct the inner surface, one needs to move the probe across the whole outer surface. This means it is moved in two dimensions implying a three-dimensional b-scan. Figure 2.15 shows a cut-out of such a b-scan where, again, each column belongs to a probe position on a line on the surface. One can see not only one but multiple reflections of the back wall and the defect. A part of the reflected ultrasound does not arrive at the probe but is reflected again at the outer surface, then reflected again at the inner surface and so on. On this way several back wall echoes with a decreasing amplitude can be measured. Another monitored signal is the surface reflection. When

the ultrasound is emitted, only a part of it enters the component, the rest is directly reflected to the probe. Thus a surface reflection can be measured right at the beginning.

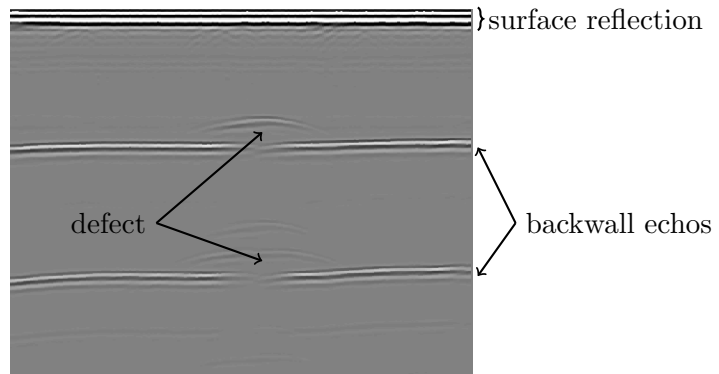


Fig. 2.15: Two-dimensional out-cut of a wall thickness measurement b-scan.

3 Defect Reconstruction: State of the Art

We have seen how ultrasonic non-destructive testing is applied in practice. Furthermore the physical background is well understood [3, 4, 7]. Thus for a known defect the ultrasonic data of a non-destructive testing method can be easily modelled. However the so-called inverse problem of reconstructing the defect out of given data, is often “ill-posed”. This means small noise on the data may lead to big errors in the reconstruction, besides there may be even no or no unique solution. To obtain numerically stable inversion algorithms one needs to use a regularisation and/or a-priori information. An introduction to inverse problems can e.g. be found in [19] and the references therein.

In the past, many heuristic and mathematical methods were developed to solve this ill-posed problem. Before we introduce our own algorithm we will consider the methods currently available in this chapter. For a better overview we classify all methods based on their original research field. The first section will present algorithms from the field of non-destructive testing. The next section is about theoretical results in inverse scattering and the last section will give an introduction to different related fields of research.

Note that although we only present the main ideas in this chapter, a complete algorithm consists of several improvements, pre- and post-processing methods. More information can be found in [20, 21] and the references in the different subsections.

3.1 Reconstruction Methods in NDT

In this section we would like to introduce two methods used in non-destructive testing. We begin with an elementary method based on gates. The second method is called “Total Focusing Method” (TFM) and includes a number of slightly different techniques as we will see. At the end of this section there will be a short subsection with an overview of other reconstruction methods being used.

3.1.1 Applying Gates

The application of a threshold [22] is a tool that is often used for denoising in signal- and image-processing. In [3, 4] a similar idea for a-scans is introduced. A gate can best be explained as a time limited threshold. It consists of a time interval and a threshold value. Depending on the type of gate, every amplitude that is measured in this time

interval and that is above (or below) the threshold value, is recorded. For example, Figure 3.1 shows a gate applied to a ToFD a-scan where the time interval begins right after the lateral wave and ends before the back wall echo. The threshold value needs to be chosen in such a way that the noise level is below the gate but a defect reflection will exceed it. Thus if we apply the gate to a whole b-scan we obtain the relevant amplitudes as shown in Figure 3.2. However, the determination of suitable threshold values is a very delicate problem depending on e.g. the noise level and probe characteristics.

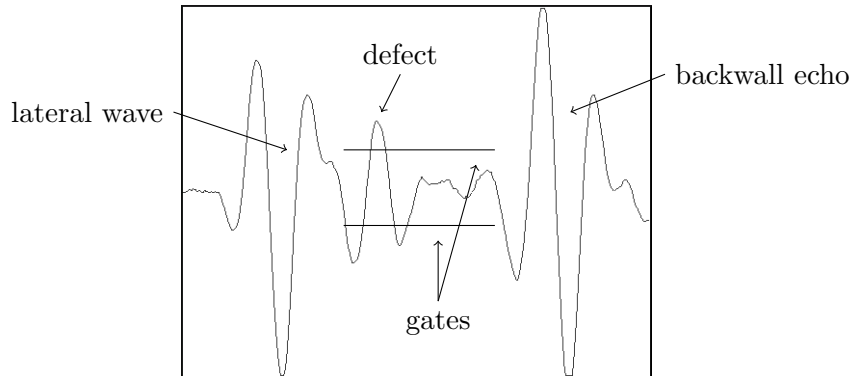


Fig. 3.1: Gate applied on ToFD a-scan.

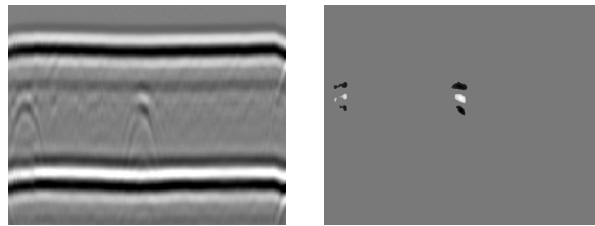


Fig. 3.2: Raw ToFD data (left) and after applying a gate (right).

Further an amplitude that falls below a certain gate can also be a very useful information. For instance if we choose the time interval such that only the back wall echo is included, an underrun of the threshold might be caused by a back wall defect [3, 4].

This technique reduces the amount of data in a fast and simple way. It can be used as a preprocessing step in automated or half-automated testing methods [23] or supportingly in manual inspections. Thus the use of multiple gates is nowadays implemented in most ultrasonic testing devices [24, 25].

3.1.2 Total Focusing Method (TFM)

The Total Focusing Method is a reconstruction method based on the amplitude arrival time similar to the methods we considered in [10]. With given ultrasound speed, probes positions and arrival time one can backtrack the signal to all possible scattering points.

For instance, if we assume the ultrasound speed to be constant, an amplitude measured at a special time can only be caused by points lying on an ellipse (or spheroid in the three dimensional case, which is an ellipsoid with two equal diameters) with the emitter and receiver positions as focal points (see Figure 3.3). For all points lying outside this ellipse the signal would be measured later and for all points inside, it would arrive earlier. Thus for all measured amplitudes one is able to find a set of points that may have caused this signal.

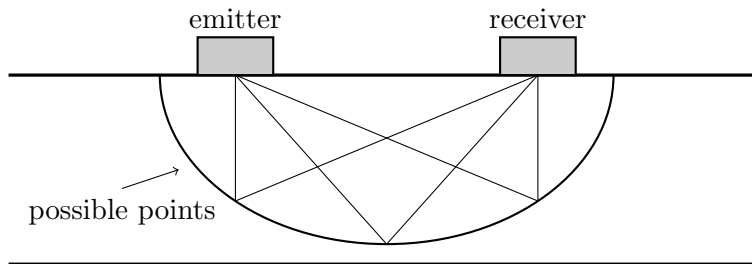


Fig. 3.3: Total Focusing Method for one emitter and receiver.

Now the reconstruction with TFM works as follows. First one discretizes the component usually by a Cartesian grid. Then for all amplitudes the set of possible scattering points is calculated and approximated with a set of points on the grid. Thus, we can assign a value to each point on the grid that is the sum of all amplitudes where this point is a possible scatterer, or 0 if the point does not belong to any set. The idea here is that the amplitudes sum up to a high value at the defect position while the value is staying low at the other grid points. Thus one can reconstruct the defect by taking the grid points with high function values.

This reconstruction technique can be used for any number of probes acting as emitter or receiver, for more information see [26]. If we use TFM with only one probe acting as emitter and receiver at once we obtain the “Synthetic Aperture Focusing Technique” (SAFT). This special case is a typical application for TFM. Detailed information to SAFT can be found in [7] or see also [3, 4, 27]. Note that the idea of TFM can also be expanded to anisotropic media [28].

Figure 3.4 illustrates the reconstruction with SAFT for a simulated ToFD dataset. The position of the material surface, the back wall and the defect is clearly noticeable. Nevertheless, TFM and SAFT do only reconstruct an image of the material. Hence, to get information about the included defect, post processing is required. Moreover, since the wave structure is still contained in the reconstruction, these methods are naturally limited in their resolution by the wave length.

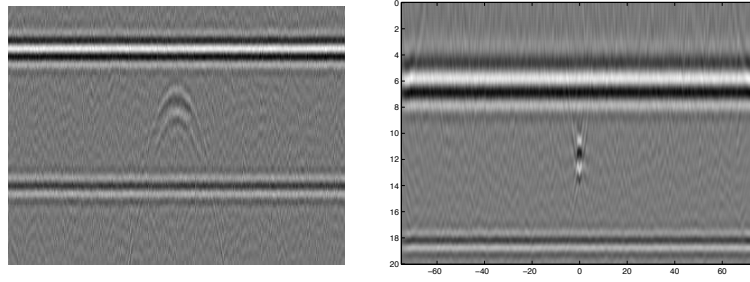


Fig. 3.4: Simulated ToFD data (left) and SAFT reconstruction(right).

3.1.3 Further Reconstruction Methods

Apart from TFM there are many other reconstruction techniques. Some of them are highly adapted to a special testing method. In this subsection we will introduce two further methods, one based on eigenvalue decomposition and one on a non-model based technique. For an overview of reconstruction methods we refer to [3, 4, 29, 30].

In [31] an algorithm using Time Reversal (TR) and Multiple-Signal-Classification (MUSIC) is presented. Here one uses the fact that the eigenvalues of the so-called Hermitian time-reversal-matrix correspond to the different scatterers. With the help of MUSIC one is able to calculate those singular values. Since the Hermitian time-reversal-matrix is a $N \times N$ matrix where N is the number of probes used, one needs at least as many probes as defects for this reconstruction technique. This method is especially used for small defects (point scatterers). For further readings also see [12, 32, 33].

Until now we have only considered reconstruction methods based on a physical model of the problem. A different approach is the use of neural networks. Depending on the input data a neural network decides whether and what kind of defect is on hand. This method possesses two disadvantages. First because of the high amount of ultrasound data the input has to be reduced to the most relevant features. Second the neural network has to be trained, meaning that one needs a set of feature vectors where the desired output is already known. With the help of this set one can adapt the network in a way so that it yields the desired output. The success of a neural network will highly depend on the choice of the features and the training set. An introduction to neural networks can be found in [34, 35], for the use in non-destructive testing see [36–38].

3.2 Inverse Scattering

Now that we have seen reconstruction methods developed in the NDT community, we will concentrate on the mathematical point of view. In applied mathematics the field of inverse scattering engages the problem of reconstructing an object out of scattered data. This can be done by solving an inverse problem involving partial differential equations

(PDEs) that describe the behaviour of waves in media. This section will give a short introduction to direct and inverse scattering and the reconstruction methods developed therein. Furthermore we will consider the limited aperture case, which discusses the reconstruction from incomplete data. For more information about inverse scattering we refer to [7, 39–42].

3.2.1 The Direct and Inverse Problem

Let us first introduce the direct and inverse scattering problem in an illustrative way for the one-dimensional case following [42]. Afterwards we will discuss the mathematical problem for the three-dimensional case based on [39].

Consider an infinite homogeneous, isotropic medium in \mathbb{R} with density p . The function $\mathbf{u}(x, t)$ with $x \in \mathbb{R}$ and $t \in [0, \infty)$ describes the displacement of the particle at position x at time t . We call \mathbf{u} the wave function. Note that, caused by oscillations, the position of a particle changes over time. Thus one normally should consider a function $\mathbf{u}(x(t), t)$ where the position is time-dependent. Assuming that the time dependence of x has only minor effects, we only study the simplified model of a linearised function $\mathbf{u}(x, t)$.

Now we assume a force that impacts this medium and analyse the behaviour of the wave function. Therefore let $F(x, t)$ be the force applied to the particle at position x at time t . Let $I = [x, x + \varepsilon]$ be an interval. Then the effective force acting on I is $F(x, t) - F(x + \varepsilon, t)$ the difference between incoming and outgoing force. The acceleration of a particle at position x is given by the second derivative of its displacement function $\mathbf{u}_{tt}(x, t)$, where we denote the derivative by a lower index. Now Newton's second law gives us

$$F(x, t) - F(x + \varepsilon, t) = p\varepsilon \cdot \frac{1}{\varepsilon} \int_I \mathbf{u}_{tt}(x, t)$$

where p is the density of the material, or in the limiting case ($\varepsilon \rightarrow 0$)

$$F_x(x, t) = p\mathbf{u}_{tt}(x, t)$$

By Hooke's law we have a linear dependence between the applied force $F(x, t)$ and the strain of the material, i.e. the change of size under this force, thus $\mathbf{u}_x(x, t)$. We obtain

$$F(x, t) = E \cdot \mathbf{u}_x(x, t) \quad \Rightarrow \quad F_x(x, t) = E \cdot \mathbf{u}_{xx}(x, t)$$

with Young's modulus E that is a material constant. Combining both formulas we obtain the wave equation

$$c^2 \mathbf{u}_{xx} - \mathbf{u}_{tt} = 0$$

with the wave velocity $c = \sqrt{E/p}$.

Now let us consider the three-dimensional case following [39]. Therefore let $x \in \mathbb{R}^3$, $t \in [0, \infty)$ and

$$\mathbf{u}(x, t) : \mathbb{R}^3 \times [0, \infty) \rightarrow \mathbb{C} \quad (3.1)$$

be the wave function. We consider $\mathbf{u} \in C^2$ to be two times continuously differentiable and $\mathbf{u}(x, \cdot) \in L^2$ for all $x \in \mathbb{R}^3$, meaning that its Fourier transform w.r.t. the time exists. Then the following notations are well-defined and the three-dimensional wave equation for an isotropic, homogeneous media is given by

$$c^2 \Delta \mathbf{u} - \mathbf{u}_{tt} = 0. \quad (3.2)$$

In inverse scattering one usually studies a slightly different equation that arises when we look at the special case of time-harmonic waves, meaning $\mathbf{u}(x, t) = \text{Re}\{\mathbf{u}(x)e^{-i\omega t}\}$ with a frequency ω . Inserting this in (3.2) and using the Fourier transform one obtains the Helmholtz equation

$$\Delta u + k^2 u = 0 \quad (3.3)$$

with $k = \omega/c$ and $u = \hat{\mathbf{u}}$ the Fourier transform of \mathbf{u} . In this section we will discuss the reconstruction of u as this is the case typically considered in inverse scattering. Later on we will see that our approach to solve the problem in time domain using the function \mathbf{u} is closely related to the theory of inverse scattering via the Fourier transform.

The general idea in scattering now is that the wave function u is a superposition of two parts $u = u^i + u^s$, where u^i is the incident wave generated by the emitters and u^s is the scattered wave. As the material defects have no influence on the incident wave u^i it should be a solution of the Helmholtz equation in \mathbb{R}^3 but does not necessarily satisfy the boundary conditions at the defect boundary. Thus u^s manipulates u in a way that the Helmholtz equation still holds outside the defect but the boundary conditions also are valid. We can formulate the scattering problem for a defect $D \subset \mathbb{R}^3$ as

$$\Delta u + k^2 u = 0 \quad \text{in } \mathbb{R}^3 \setminus D \quad (3.4)$$

$$u = u^i + u^s \quad (3.5)$$

$$u = 0 \quad \text{on } \partial D \quad (3.6)$$

$$\lim_{r \rightarrow \infty} r \left(\frac{\partial u^s}{\partial r} - iku^s \right) = 0, \quad (3.7)$$

with $r = \|x\|_2$. Condition (3.7) is the so-called *Sommerfeld radiation condition*. For every outgoing scattering wave there is always an incoming wave as solution which is physically not relevant but does not hold the Sommerfeld radiation condition. Thus, this condition guarantees outgoing scatterers. The boundary condition (3.6) is called

Dirichlet condition and implies a sound-soft medium. It can be replaced by the Neumann condition

$$\frac{\partial u}{\partial v} = 0 \text{ on } \partial D \quad (3.8)$$

for sound-hard media, where v is the outer normal on ∂D , or also other boundary conditions as *impedance boundary condition*, *transmission condition* or *resistive boundary condition*, see [39, 43]. Finally one can replace (3.4) and (3.6) by

$$\Delta u + k^2 n(x)u = 0 \text{ in } \mathbb{R}^3 \quad (3.9)$$

and obtains the scattering problem for the more general case of inhomogeneous media described by the function $n(x)$.

Now let us introduce the far field of a function u . It was shown in [39] that every scattered wave that satisfies the Sommerfeld radiation condition can be written as

$$u^s(x) = \frac{e^{ik\|x\|_2}}{\|x\|_2} \left[u_\infty \left(\frac{x}{\|x\|_2} \right) + \mathcal{O} \left(\frac{1}{\|x\|_2} \right) \right], \quad \|x\|_2 \rightarrow \infty, \quad (3.10)$$

where u_∞ is supported on $\mathcal{S} := \{x \in \mathbb{R}^3 \mid \|x\|_2 = 1\}$. We call u_∞ the *far field* of u . In many inversion algorithms one assumes that the measured data is approximative the far field u_∞ . Thus, one does not require measurements covering \mathbb{R}^3 but only $r \cdot \mathcal{S}$ for a fixed $r \in (0, \infty)$. One can show that the far field satisfies

$$u_\infty(x) = \frac{1}{4\pi} \int_{\partial D} \left[u(y) \frac{\partial e^{-ikx \cdot y}}{\partial v} - \frac{\partial u}{\partial v}(y) e^{-ikx \cdot y} \right] dy, \quad x \in \mathcal{S}, \quad (3.11)$$

with $x \cdot y$ denoting the standard scalar product in \mathbb{R}^3 and v being the outer normal on ∂D .

Now the direct scattering problem can be formulated as follows. For given u^i and D or $n(x)$ construct the far field u_∞ . The inverse scattering problem is to reconstruct D or $n(x)$ for given u^i and u_∞ .

At this point, we will turn our attention to the reconstruction techniques developed in inverse scattering. We will return to the theory in Section 4.3 where we discuss the analogy between our model and the inverse scattering theory. Indeed, we will show that our inversion method first introduced by a heuristic approach in this work, can also be evolved from the theory of inverse scattering. Thus we find ourselves right in the middle, building a connection between the practical and the theoretical world. Furthermore, we will see that our algorithm possesses some similarities with SAFT, which, first a heuristic approach, was later also derived out of the inverse scattering theory [7, 44].

3.2.2 Reconstruction Methods for Complete Data

We will first study reconstruction methods considering the best case of completely given data, meaning u^s or u_∞ are completely known. While we only present a few algorithms and their ideas, a wide overview can be found in [39, 45–47] and the references therein. A short introduction can also be found in [48]. The following outline is based on the references mentioned above.

Basically, almost every reconstruction method can be assigned to one of the three classes, namely iterative methods, decomposition methods and sampling methods (sometimes called probe methods). We will provide the general ideas of these three classes and give an illustrative exemplary algorithm.

We will start with the class of iterative methods, which is based on the most straightforward idea. The inverse scattering problem for a fixed incident wave u^i can be expressed by an operator F such that

$$F(\partial D) = u_\infty.$$

Given the far field u_∞ one wants to find the defect boundary ∂D , meaning that one needs to invert the non-linear, ill-posed operator F . We may apply the known theory for non-linear inverse problems [19] and make use of the methods developed therein which often are iterative. This forthright approach leads to regularised reconstruction methods based for example on Newton- or Landweber-Iterations [45]. While most of these techniques try to linearise the non-linear operator locally, the *Fréchet derivative* of F is required, which is a crucial step in these algorithms, see [19, 49].

To give a demonstration of such algorithms, we will comprehend an example given in [45]. Here only star shaped defects are considered, such that ∂D can be expressed by a function $r : \mathcal{S} \rightarrow (0, \infty)$ in the sense that $\partial D = \{r(x)x \mid x \in \mathcal{S}\}$. Now we define $F(r) := F(\partial D) = u_\infty$. Considering two approximations D_1 and D_2 of the defect D expressed by functions r_1 and r_2 we get the linearisation of F around D_1 by

$$F(\partial D_2) = F(r_2) = F(r_1 + (r_2 - r_1)) \approx F(r_1) + F'(r_1)(r_2 - r_1),$$

where $F'(r_1)$ is the Fréchet derivative of F at r_1 . For a given far field u_∞ starting with a function r_0 we can use a regularised Newton iteration to solve the inverse problem. This means, starting with $k = 0$, iterate

Solve: $F(r_k) + F'(r_k)r = u_\infty,$

Update: $r_{k+1} = \lambda r_k + (1 - \lambda)r$ and $k = k + 1,$

until a stopping criterion holds, where $\lambda \in (0, 1)$ is a regularisation parameter. Such iterative approaches typically lead to accurate and stable reconstructions. However, the disadvantages of such methods are obvious. A good and stable implementation of both F and F' is required. Furthermore, as a Newton method only converges to local minima

one needs a good starting guess, i.e. good a-priori information. For a detailed numerical discussion we refer to [50–52].

The second class of methods is the class of decomposition methods. As we have seen before, the inverse scattering problem is non-linear and ill-posed. Decomposition methods try to deal with this problem by separating it into two parts, one being ill-posed but linear and the other non-linear part. Thus, one can deal with the two difficulties separately. A common way to do so is to reconstruct u^s out of u_∞ first and then reconstruct ∂D out of u^s .

$$u_\infty \xrightarrow[\text{ill-posed}]{\text{linear}} u^s \xrightarrow{\text{non-linear}} \partial D$$

As an example for this kind of techniques we will outline a method by Kirsch and Kress that can be found in [39, 45, 47]. The authors show that for an unknown density φ in the interior of D , there is a linear but ill-posed integral operator F_1 such that

$$F_1\varphi = u_\infty$$

holds. For the same density φ one can define an integral operator F_2 such that u^s is given by

$$F_2\varphi = u^s.$$

Now for a known incident wave u^i one can find ∂D by analysing where $u^i + u^s$ satisfies the boundary condition. For Dirichlet boundary conditions we obtain the algorithm:

- Solve** the inverse problem $F_1\varphi = u_\infty$ by regularisation.
- Calculate** $u^s = F_2\varphi$.
- Solve** the non-linear problem $u^i + u^s = 0$ on ∂D .

A different formulation of this idea as a minimisation problem in φ and ∂D can be found in [39]. Here one tries to solve both problems at once in a least square sense by minimizing the functional

$$\|F_1\varphi - u_\infty\|_{L^2(\mathcal{S})}^2 + \alpha\|\varphi\|_{L^2}^2 + \gamma\|u^i + F_2\varphi\|_{L^2(\partial D)}^2$$

over φ and ∂D , where α is a regularisation parameter and γ is a coupling parameter. Another decomposition method is the *point-source method* introduced by Potthast [53, 54]. We will later discuss this method for the case of limited data. While the numerics of iterative methods are slightly better, this type of method does not require an implementation of the forward solver and its Fréchet derivative.

The last class of methods we would like to consider are sampling methods. The idea here is to define a functional $f(x)$ that decides whether a point $x \in \mathbb{R}^3$ is on the boundary ∂D or not. In practice $f(x)$ will be defined in a way that it tends to infinity when x

tends to a point on the boundary. If one takes a finite grid $G \subset \mathbb{R}^3$ of the relevant area in \mathbb{R}^3 one can approximate ∂D by

$$\partial D = \{x \in G \mid |f(x)| > K\}$$

with a constant $K > 0$. Note that this reconstruction idea is similar to the idea of the TFM/SAFT algorithm. An example for such an algorithm is the linear sampling method given in [45, 55]. We define the operator F as

$$Fg(x) = \int_{\mathcal{S}} u_{\infty}(x, s)g(s)ds, \quad x \in \mathcal{S}$$

where $g \in L^2(\mathcal{S})$ and $u_{\infty}(\cdot, s)$ is the far field for an incident plane wave with direction s , i.e. $u_i(x, s) = e^{ikx \cdot s}$. Now one can show that the solution $g(x, z)$ of

$$Fg(\cdot, z) = \Phi_{\infty}(\cdot, z) \quad \forall z \in \mathbb{R}^3$$

tends to infinity if z approaches the boundary of D , meaning $\|g(\cdot, z)\|_{L^2(\mathcal{S})} \rightarrow \infty$. Here $\Phi_{\infty}(x, z) = \frac{1}{4\pi}e^{-ikz \cdot x}$ is the far field of the fundamental solution $\Phi(x, z) = \frac{e^{ik\|x-y\|_2}}{4\pi\|x-y\|_2}$, see [39, 45]. With $f(z) = \|g(\cdot, z)\|_{L^2(\mathcal{S})}$ we obtain the algorithm

Solve $Fg(\cdot, z) = \Phi_{\infty}(\cdot, z)$ for all $z \in \mathbb{R}^3$.

Set $\partial D = \{z \in \mathbb{R}^3 \mid \|g(\cdot, z)\|_{L^2(\mathcal{S})} > K\}$ for a constant $K > 0$.

One major advantage of these methods is that they do not require the knowledge of a-priori information or the boundary condition. However, a computation of g for all $z \in \mathbb{R}^3$ is required and the accuracy of the approximation highly depends on the choice of K .

Now we have demonstrated different ideas and algorithms to solve the inverse scattering problem. Although we have restricted our introduction to the case of homogeneous data, there are of course similar reconstruction ideas for the inhomogeneous case, see e.g. [56, 57]. However, we only considered algorithms for complete data. In our application of non-destructive testing often data measured only from a limited number of directions is given. In addition, because of the measurement set-up, it is also not clear if the measured data approximates the far field well enough. This means that although the theory shown above provides good reconstruction methods for complete data, we cannot apply it to our problem. Hence we now move on to the limited aperture case, which will supply reconstruction methods for the case of incomplete data.

3.2.3 The Limited Aperture Case

Let Γ be a proper subset of \mathcal{S} . Then the reconstruction problem in the limited aperture case is to reconstruct the scatterer D when u_{∞} is only known on Γ . Although this

problem is more difficult than the full aperture case, analytically it is sufficient to know u_∞ on a subset Γ with a non-empty interior to provide uniqueness of the problem, see [39]. However, in practice the stability and quality of the reconstruction decreases drastically especially if Γ is small [39]. For this reason, good reconstruction methods are rare and are often restricted to special cases and a-priori assumptions. We will shortly present three methods published in the last years that try to transfer the reconstruction ideas of the full aperture case. Their use for our application will be discussed. Note that we will skip the analysis of the algorithms and only introduce the ideas, since examples have already been shown in the last section. More information about the presented methods can be found in [58–62].

We start with the point source method for the limited aperture case presented in [58, 59]. This method belongs to the class of decomposition methods and especially considers incident plane waves. The authors define two operators $H : L^2(\Gamma) \rightarrow L^\infty(\mathbb{R}^3)$ and $A_g : L^2(\Gamma) \rightarrow L^2(\mathbb{R}^3)$, where A_g depends on a density g defined on Γ . Now one can show that if Hg satisfies some properties, then $A_g u_\infty$ is a good approximation of u^s at least on a special subset Ω of \mathbb{R}^3 . This subset Ω is independent from the incidence direction and frequency but it is unknown. The idea of Potthast and Luke is to use the point source method for several incident plane waves with different directions and/or frequencies. One can now detect Ω by a combination of all reconstructions, since their behaviour differs outside Ω but is equal inside Ω . However, although the presented method can reconstruct the defect quite well even with few data, it still requires different frequencies and incident directions. The numerical results in [58] show a breakdown of this algorithm for very small apertures. It is therefore too unstable for our application.

Next, we discuss a method presented in [60] for the special case of photon scanning tunnelling microscopy. It is an iterative method where the Born approximation is used for a good starting guess. Especially for low frequencies this is a good approximation. This allows an iterative reconstruction of the defect with multiple frequency data. Starting with the Born approximation for the lowest frequency one makes use of a connection between two solutions of the inverse scattering problem with slightly different frequencies to obtain an update step. Numerical experiments show a quite good reconstruction, but the use of several frequencies is crucial, thus an adaption to our problem is not possible.

Finally, we mention the enclosure method introduced in [61]. The enclosure method tries to reconstruct the convex hull of an obstacle. Note that [61] is restricted to the two-dimensional case. The algorithm uses the fact that the convex hull can be reconstructed from the knowledge about the function

$$h_D(\omega) := \sup_{x \in D} x \cdot \omega$$

with $\omega \in \mathcal{S}$, where in this case \mathcal{S} denotes the two-dimensional sphere. To calculate the function values of h_D one first needs to construct an approximate density g similarly to the point source method. Then one can define an operator $I_g(\omega, r)$ based on the incomplete data u_∞ such that I_g is approximately linear in r with slope $h_D(\omega)$. Thus,

calculating $I_g(\omega, r)$ for fixed ω at several points r_i , $i = 1, 2, 3$ one can fit a line through the points $(r_i, I_g(\omega, r_i))$ and take its slope as an approximation for $h_D(\omega)$. For each calculated value of h_D the convex hull can be reduced to a half plane defined by ω . That way one can restrict the convex hull of D iteratively to a shrinking area. One advantage of this method is that it does not need multiple frequencies or incident directions. Nevertheless, the method cannot reconstruct non convex structures. This specifically applies for cases with more than one scatterer since in this case the area between the obstacles will be included in the convex hull as one can see in the numerical results in [61]. Furthermore the reconstructed area even for the noiseless case contains several wrongly reconstructed parts.

As we have seen, there were several ideas to solve the inverse scattering problem in the case of limited aperture. However, most algorithms try to compensate the missing data by using several frequencies or incident waves. Such data is often not available in non-destructive testing. For a reconstruction with a very small amount of data the theory is limited and the provided algorithms are unstable and/or improper.

3.3 Other Fields of Research

In the last section of this chapter we would like to take a look at research fields that are related to non-destructive testing. As there are many similar problems, the reconstruction algorithms used might be helpful for our case.

Let us begin with a summary of the radar technique based on [63,64]. Radars have to detect obstacles (e.g. air planes) in an approximately homogeneous medium (e.g. air) using scattered signals from high frequency waves. So the detection of obstacles with radar beams is similar to the problem of defect reconstruction in NDT. Here one uses two different approaches. The first one is to upgrade the radar aperture in a way that the transmitted waves have a very special form. Thus the amplitude of the reflected waves corresponds directly to the direction of the scatterer while the distance to the object can be calculated using the runtime of the signal. As a second approach one uses algorithms to solve the inverse scattering problem. Here, one inversion method has proven to deliver good results. The *Synthetic Aperture Radar* (SAR) is a technique to combine radar signals taken at different positions. The affinity between the names SAR and SAFT is not a coincidence, indeed both methods are based on the same idea applied in different fields of research [30]. Yet another implementation of this idea is used in sonar techniques, which uses the same modelling principles as the radar technique but underwater. The corresponding algorithm is called *Synthetic Aperture Sonar* (SAS) [65].

Another large field of application for ultrasound can be found in medical or biomedical imaging. First it seems that scanning a human body or a steel tube for “defects” should work similarly and thus the same algorithms should be applied. But there are big differences between these two applications: In particular, there are different requirements

to the reconstruction results. In non-destructive testing, the reconstruction only needs to give a rough impression about the enclosed defect, but the algorithm has to deal with strong noise and needs to be fast (e.g. for ToFD a tube with length 0.5m needs to be inspected in 1s). The opposite case applies to medical applications. Here it is crucial to have a detailed reconstruction of e.g. brain tumours. Therefore one accepts methods that have a low noise level but may take a lot of time. Thus, one usually has enough data to apply the inversion theory known from inverse scattering. This is usually not the case for NDT. For more details about medical and biomedical imaging we refer to [12,66,67].

As we have seen, the methods and theory in all these different applications are largely consistent. While in our application the inverse scattering theory for strongly limited data is just at the beginning, the heuristic inversion approach with TFM/SAFT currently seems to be the method of choice. In the next chapters we will introduce our new model and a suitable inversion algorithm. But we keep in mind that in the end we have to compare our numerical results with the TFM/SAFT algorithm.

4 Model

In the previous chapters we have gathered information about non-destructive testing methods, their historical and physical background and the known reconstruction methods. We have seen that there are only few applicable reconstruction algorithms due to the limited amount of data and the high noise level. Furthermore, the methods used today still have drawbacks. For this reason, we now develop a new model based reconstruction method. We introduce it step by step in the following chapters.

We start by introducing our new model for ultrasonic testing methods. The first section describes the continuous model which is mainly based on a geometrical approach. At the end of the section we add a convolution to the model to include the wave structure of ultrasound. After that we have a look at the discretization. The last section of this chapter gives a validation, first by a comparison between modelled and real data, and secondly by the theory of inverse scattering that we have seen in the previous chapter. Here we also analyse the type of noise in real data and add it to our model. As we will see, our model is strongly related to inverse scattering, therefore our model function will be denoted by u to fit into the notation we have used so far.

4.1 The Continuous Model

In this section we design our continuous model for ultrasonic testing. Thereby this model has to fulfil several requirements. On the one hand, it should deliver a good approximation for real data, meaning the forward problem should be approximated accurately. On the other hand we want to keep the model as simple as possible to be able to develop a fast and stable reconstruction method later on that can deal with the limited amount of data. Furthermore, the model should be applied to different testing set-ups. This means, it needs to be flexible relating to e.g. material and probe properties and positions.

Taking all this into account we shall develop a new model based on simple geometrical relations combined with a convolution. Please note that we therefore use fundamental facts from geometrical optics, optical physics or its applications, see e.g. [3,4,7,11,68–70]. However, to the author's best knowledge, a model such as the one we will present in this work, does not exist so far.

In our model we will only consider one emitter and one receiver. However, to simulate a measurement based on multiple probes one can simply use this model for every com-

combination of emitters and receivers. Furthermore, we will neglect the lateral wave as it is not relevant for the reconstruction and can be easily filtered out. Effects caused by the material boundary e.g. back wall echo and surface reflections will often be skipped in the simulations but can be included by adding the material boundary as a “defect” to the simulation.

4.1.1 The Geometrical Approach

We start this subsection by introducing the functions and variables we need to describe our model. Consider a measurement set-up with one emitter and one receiver probe on a homogeneous infinite material with several defects inside. The probes are considered to be idealised points. Now we use the following notations.

Denote by $x_E(h) : [0, 1] \rightarrow \mathbb{R}^3$ and $x_R(h) : [0, 1] \rightarrow \mathbb{R}^3$ the path of the emitter and receiver probe, meaning both probes move along a path in \mathbb{R}^3 starting at $x_E(0)$ respectively $x_R(0)$ and ending at $x_E(1)$ respectively $x_R(1)$. For each $h \in [0, 1]$ a measurement (an a-scan) is done with the probes positions $x_E(h)$ and $x_R(h)$. Note that while these functions imply a one-dimensional ordering of the a-scan by the parameter h , for some testing methods like wall thickness measurement it is more reasonable to consider the data in more dimensions. This can be done easily by rearranging the data afterwards. In Figure 4.1 we demonstrate an example where the a-scans are arranged on a two-dimensional grid.

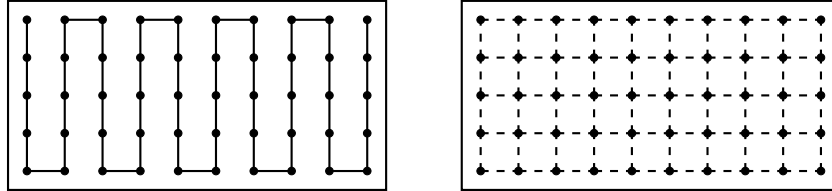


Fig. 4.1: Way of the probes (left) and arrangement grid (right).

Next, let $K \in \mathbb{N}$ be the number of defects and let the defects be given by the disjoint sets $D_k \subset \mathbb{R}^3$ for $k \leq K$, where each set is connected but $D_k \cup D_j$ is not connected for $k \neq j$.

Furthermore we denote the geometrical model function by $\tilde{u}(h, t) : [0, 1] \times [0, \infty) \rightarrow \mathbb{R}$, where $\tilde{u}(h, t)$ is the measured amplitude (of the geometrical model) at time t with probes positions at $x_E(h)$ and $x_R(h)$. Here $t = 0$ is the time, where the emitter sends out an ultrasound impulse at position $x_E(h)$.

The goal of this subsection is to determine the geometry of the measurement set-up by describing \tilde{u} with the help of x_E , x_R and $\{D_k\}_{k=1}^K$. Therefore we treat the ultrasound as a ray and ignore its wave structure, or in other words, we assume the ultrasound impulse to be a peak function.

As mentioned above, $\tilde{u}(h, t)$ models the amplitude at time t for probe positions $x_E(h)$ and $x_R(h)$. To construct the function \tilde{u} we have to answer two questions:

When will an amplitude be measured at the receiver, i.e. for which $h \in [0, 1]$ and $t \geq 0$ is $\tilde{u}(h, t) \neq 0$?

What determines the amplitudes strength, i.e. what determines the value $\tilde{u}(h, t)$?

Let us have a closer look at the first question. The ultrasound has to travel through the material before it reaches the receiver. Starting at the emitter position $x_E(h)$ it moves towards the defect. There it is reflected/diffracted into the direction of the receiver (see Figure 4.2 for e.g. reflection).

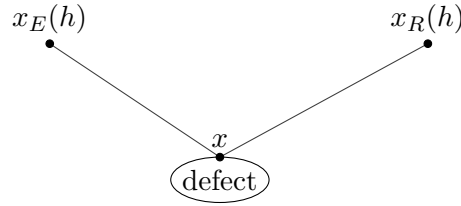


Fig. 4.2: Sketch of the geometry for a reflecting point x .

Let $x \in \mathbb{R}^3$ be the point on the defect boundary where the ultrasound is reflected or diffracted. Then the distance the ultrasound has covered is given by $\|x - x_E(h)\|_2 + \|x_R(h) - x\|_2$. We assume homogeneous material, thus the ultrasound speed c is constant. Given that the ultrasound was send out at time $t = 0$ the amplitude will arrive at the receiver at time

$$\frac{\|x - x_E(h)\|_2 + \|x_R(h) - x\|_2}{c}.$$

This means the amplitude of the reflection/diffraction caused by x will influence the value $\tilde{u}(h, t)$ if and only if

$$t = \frac{\|x - x_E(h)\|_2 + \|x_R(h) - x\|_2}{c} \tag{4.1}$$

holds. Vice versa, to answer our first question, the value $\tilde{u}(h, t)$ can only be unequal 0 if there exists at least one point $x \in \cup_{k=1}^K \partial D_k$ such that (4.1) holds. Here ∂D_k denotes the boundary of the defect D_k . With the help of the next definition we can restate this result.

Definition 4.1.1:

Let $E_{h,t} \subset \mathbb{R}^3$ be the set of all points for which (4.1) holds, i.e.

$$x \in E_{h,t} \quad \Leftrightarrow \quad t = \frac{\|x - x_E(h)\|_2 + \|x_R(h) - x\|_2}{c}.$$

Note that $E_{h,t}$ is a spheroid with semi-major-axis $tc/2$ and focal points $x_E(h)$ and $x_R(h)$.

We directly obtain:

Model Assumption 4.1.2:

If $\tilde{u}(h, t) \neq 0$ for $h \in [0, 1]$ and $t \geq 0$ then

$$E_{h,t} \cap \left(\bigcup_{k=1}^K \partial D_k \right) \neq \emptyset. \quad (4.2)$$

Note that the contrary statement does not have to be true. This phenomenon is called *destructive interference* [5]. Here several ultrasound waves are incoming at the receiver but their (positive and negative) amplitudes add up to 0 and thus $\tilde{u}(h, t) = 0$.

Given this assumption, we can now consider by what the amplitude is influenced if (4.2) holds. Let us follow the way of the ultrasound from emitter to receiver and debate its effects on the amplitude.

The ultrasound wave is generated at the emitter and hence surely depends on the emitter's properties. For example the piezoelectric effect (see Subsection 2.2.1) indicates that a high voltage generates waves with large amplitudes and a low voltage generates small waves. But nowadays probes go far beyond this and can even create ultrasonic waves whose amplitude depend on the direction. Thus it is possible to e.g. focus on special directions. We use the function $\mathbf{a}(s) : \mathcal{S} \rightarrow \mathbb{R}$ to simulate the emitter's behaviour, where \mathcal{S} denotes the unit sphere. It maps a direction to the corresponding amplitude. The function \mathbf{a} strongly depends on the emitter type and can be determined with a special measurement process see e.g. [3, 4, 8]. Figure 4.3 shows an example for such an emitter function where the amplitudes are colour-coded onto the sphere. One can see the side lobes that are typical for a focussing emitter. Note that the direction s from $x_E(h)$ to a point x can be derived via $s = \frac{x_E(h) - x}{\|x_E(h) - x\|_2}$.

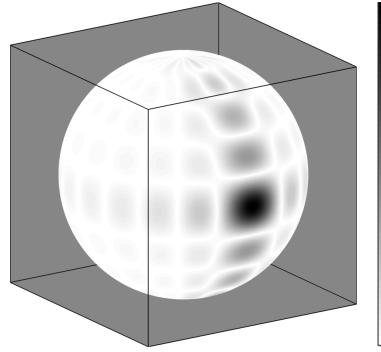


Fig. 4.3: Focussing directional amplitudes.

After the ultrasound wave is generated it begins to move through the object. This movement will constantly change the amplitude by many effects often grouped under attenuation [3–5, 8]. Here we will only consider three sources of attenuation to keep the model simple: Divergence, absorption and dispersion. Neglecting other sources will

generally not worsen the model as these three sources are usually the main reasons for attenuation.

Let us start by considering divergence, see also [6, 71]. As we assumed homogeneous and isotropic media, the ultrasound will spread out from its source in each direction with the same speed. Thus, the wave front will form a sphere with radius ct where t is the elapsed time. This means, the wave front area is increasing in time. Thus, to keep the carried energy constant, the amplitudes have to decrease. Let us have a look at the energy on a wave front at time t , where \tilde{a} is a spherical function of amplitudes at time t . As the energy is proportional to the squared amplitude [71], we obtain

$$\int_{tcS} \tilde{a}^2\left(\frac{x}{tc}\right) dx = (tc)^2 \int_S \tilde{a}^2(x) dx \stackrel{!}{=} \int_S a^2(x) dx,$$

where the last equation has to hold to keep the energy constant [6]. This holds, if $\tilde{a} = \frac{a}{tc}$, thus all amplitudes have to decrease in time inversely proportional to the sphere radius tc . In other words, the greater the wave front gets the smaller the amplitude will be at each point. Figure 4.4 shows the increase of a spherical sector in time. Note that the decrease factor is only true for spherical waves, for other wave forms one has to estimate the factor similarly.

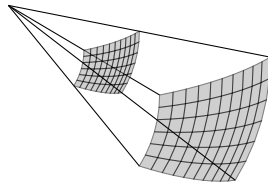


Fig. 4.4: Spherical increase in time.

Let us now consider the effects of absorption and dispersion, see also [3, 4, 72]. An ultrasonic wave is the motion of particles which causes heat e.g. by friction. This heat is a conversation of energy, thus reduces the amplitude of the wave. It seems to be clear that the more motion the wave causes (i.e. the higher the frequency or amplitude is) the more energy will be lost by heat. A reasonable and often used conclusion is that the loss of energy by absorption is proportional to the amplitude [3–5, 8, 72]. Thus one obtains an exponential decay. The effect of dispersion is quite similar. Although we have assumed that the material is homogeneous, there are tiny inhomogeneities smaller than the wave length that can affect the ultrasound. These inhomogeneities are too small to be considered as scattering objects, but nevertheless they cause a dispersion of the wave. This means, while most of the wave will pass the inhomogeneity without being influenced much, part of it will be spread in all directions. Again we obtain a decrease proportional to the amplitude. Because the inhomogeneities are evenly distributed, we

get an exponential decrease by the dispersion. Altogether one can summarize that the amplitude decreases by a factor of

$$e^{-\gamma t}$$

where $\gamma > 0$ is a constant depending on the material and the frequency of the wave. For a list of values for several materials and frequency we refer to [3, 4, 8]. Note that besides the decrease of the amplitude, dispersion has another effect on the measured data. The part of the ultrasound that is spread over and over in the material leads to a significant noise on the data as it is also measured at the receiver. We will consider this type of noise later in Section 4.3.

Now we know how the amplitude changes while the ultrasound moves through the material. The next point we want to consider, is the change of the amplitude when the wave “hits” a surface, i.e. the influence by reflection or diffraction. We start with the diffraction as it is simpler to analyse. We remember that, due to Huygens principle for diffraction (see Figure 2.7), the wave does not interact with the boundary but moves alongside it until the surface ends and the diffracted wave accrues. Thus, the diffracted wave has the same amplitude as the incoming wave as long as we do not consider interference with other waves. Note that there are several models for diffraction that consider interference as for example Fresnel diffraction or Fraunhofer diffraction (see [69]). For our model, it is enough to consider interference not before it arises at the receivers position. This means we will treat two interfering waves separately until they reach the receiver (see later discussions).

Let us now have a look at the change of amplitude caused by reflection. Differently to diffraction, here the wave interacts with the surface and part of it penetrates it as a refracted wave. This means, a part of the energy is split up and lost what leads to a decrease of the amplitude. This is a complicated process influenced by many different factors. The reflected amplitude does not only depend on the angle of incidents but although on the wave types of incident and reflected wave, the state (solid, fluid, gaseous) and density of both bordering materials and the speed of both longitudinal and transversal wave. Here we only introduce the formula for the most important case of an incoming longitudinal wave reflected as a longitudinal wave at a border of solid (steel) and gaseous material (gas inclusions). The amplitude decrease is then given by

$$\mathfrak{Ref}(\alpha) := \frac{\left(\frac{c_t}{c_l}\right)^2 \sin(2\alpha) \sin(2\tilde{\alpha}) - \cos^2(2\tilde{\alpha})}{\left(\frac{c_t}{c_l}\right)^2 \sin(2\alpha) \sin(2\tilde{\alpha}) + \cos^2(2\tilde{\alpha})}, \quad (4.3)$$

where the ultrasound speed of longitudinal waves $c_l = c$ has an explicit index to distinguish it from the speed of transversal waves c_t . Furthermore, $\tilde{\alpha}$ is given by Snell’s law $\sin \alpha / \sin \tilde{\alpha} = c_l / c_t$ with the incident angle α . To ensure that the incident angle is well defined, we have to assume that the boundaries of all defects D_k are regular. Then the

outer-pointing normal $N_k : \partial D_k \rightarrow \mathcal{S}$ on D_k exists and one can determine the angle of incident for a point $x \in D_k$ and emitter position $x_E(h)$ by

$$\cos \alpha = \left\langle \frac{x_E(h) - x}{\|x_E(h) - x\|_2}, N_k(x) \right\rangle. \quad (4.4)$$

For the detailed mathematics and formulas for all cases we refer to [73]. An overview can also be found in [3,4]. Figure 4.5 shows the absolute decrease of the amplitudes depending on the incident angle for the case of a gas inclusion in steel with $c_l = c = 5.94 \frac{\text{mm}}{\mu\text{s}}$ and $c_t = 3.29 \frac{\text{mm}}{\mu\text{s}}$. Note that the function in (4.3) is negative for all $\alpha \in [0, \pi]$ causing a phase shift of the reflected wave. Figure 4.5 shows the absolute value of $\Re f(\alpha)$ in 4.3.

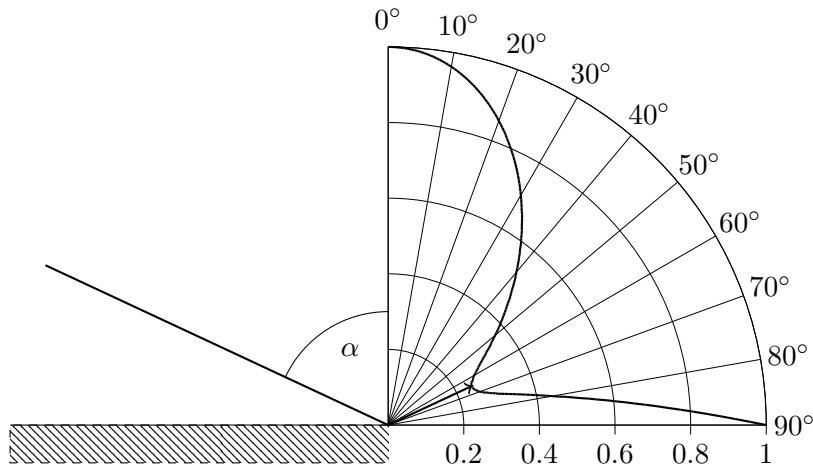


Fig. 4.5: Amplitude of reflected wave depending on incident angle.

So far we have only considered principles of ray optics. According to ray optics, the angle of incident is the same as the emergent angle. However, the direction from the defect point x to the receiver $x_R(h)$ may differ from the emergent direction by an angle θ . Using only ray optics means the reflected wave will miss the receiver and will never be recorded. Figure 4.6 (top) illustrates this problem. However, we know that the reflected wave will propagate not only in the emergent direction but will also be spread along others. Thus, part of it will reach the receiver but with a decreased amplitude depending on the angle θ . To simulate the behaviour of a reflected very thin ray-like ultrasound wave we use a simple physical consideration, the so called *diffraction at a circular aperture* [6,68]. Thereby a plane wave is transmitted onto a non penetrable object that has a circular hole with diameter d . The wave can only pass the object through the circular opening. This means, for small d only a very thin ultrasound wave will pass it and diffract in all directions on the other side of the object. To solve our ray-optics problem we consider a circular aperture with radius d that is placed in emergent direction with distance \tilde{d} to the scatterer x (see Figure 4.6 (bottom)). The ultrasound will then diffract towards the receiver with an angle $\tilde{\theta}$. If we now let \tilde{d} go to 0, we obtain $\tilde{\theta} = \theta$ and the amplitude

of the ultrasound wave which propagates into receiver direction can be described as a function of θ by the diffraction at a circular aperture [6, 68]

$$\frac{J_1\left(\frac{\pi d}{\lambda} \sin \theta\right)}{\frac{\pi d}{2\lambda} \sin \theta}, \quad (4.5)$$

where the Bessel-function of the first kind J_1 [39] is given by

$$J_1(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(k+1)!} \left(\frac{x}{2}\right)^{2k+1}.$$

This consideration is only reasonable if $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. For $|\theta| > \frac{\pi}{2}$ we will set the reflected amplitude to 0 as the receiver's direction differs strongly from the reflected direction (see also Figure 4.8). The angle θ can be derived according to Figure 4.6 (top) by

$$\cos \theta = \left\langle \frac{x_R(h) - x}{\|x_R(h) - x\|_2}, 2 \left\langle \frac{x_E(h) - x}{\|x_E(h) - x\|_2}, N_k(x) \right\rangle N_k(x) - \frac{x_E(h) - x}{\|x_E(h) - x\|_2} \right\rangle. \quad (4.6)$$

Note that, to simulate an *ultrasound ray* with thickness 0 one may also consider the case $d \rightarrow 0$. This will lead to the scattering at a point scatterer x , i.e. the point x will have no outer surface normal as it has no surface at all and thus the amplitude will be the same in all directions. So instead of considering the case $d \rightarrow 0$, we keep d as a variable of our function as this approach will prove to be very useful in the discretization of our model. The function in (4.5) is shown in Figure 4.7 with $\frac{d}{\lambda} = 4$.

Knowing how the amplitude is changed by reflection and diffraction, we need a criterion to determine whether a signal is reflected or diffracted at a point x . This can easily be done with the help of the incident angle $\alpha \in [-\pi, \pi)$. We note that an ultrasound wave will be reflected at a point x if the incident angle satisfies $|\alpha| < \frac{\pi}{2}$. A diffraction will only appear if the ultrasound wave approaches the point x with an incident angle $|\alpha| \geq \frac{\pi}{2}$. Figure 4.8 illustrates the angles of incident for reflection and diffraction at a point x .

Following the considerations above and the illustrations of Figure 4.8 we now define

Definition 4.1.3:

Let all sets ∂D_k be regular. Then the *reflection-diffraction-term* $\mathfrak{RD} : [-\pi, \pi) \rightarrow \mathbb{R}$ is defined by

$$\mathfrak{RD}(\alpha) := \begin{cases} \Re\{f(\alpha)\} \frac{J_1\left(\frac{\pi d}{\lambda} \sin \theta\right)}{\frac{\pi d}{2\lambda} \sin \theta} & |\alpha| < \frac{\pi}{2}, |\theta| \leq \frac{\pi}{2} \\ 0 & |\alpha| < \frac{\pi}{2}, |\theta| > \frac{\pi}{2} \\ 1 & |\alpha| \geq \frac{\pi}{2} \end{cases}$$

where the angle θ can be derived by (4.6).

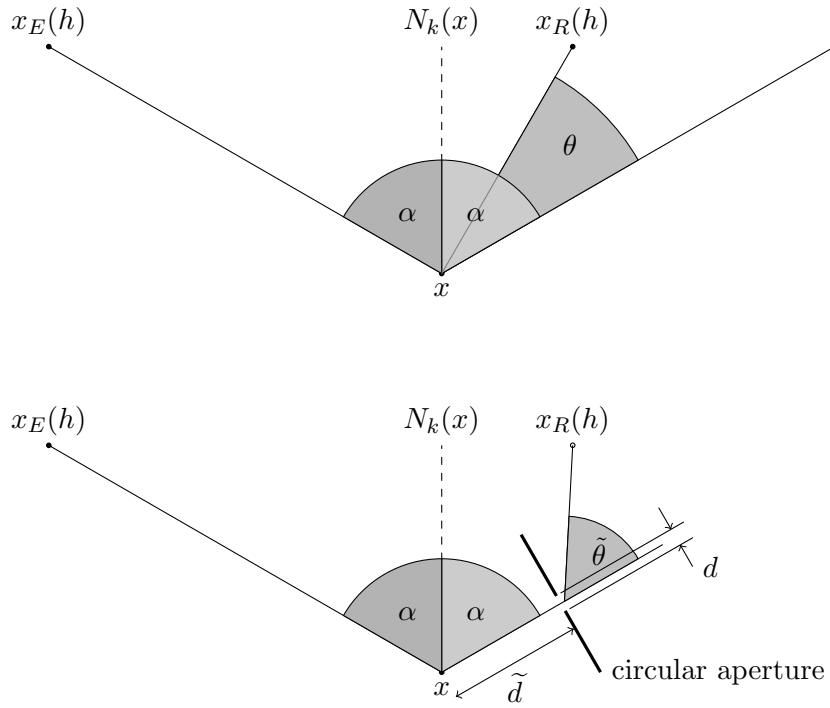


Fig. 4.6: Diffusion approximation with diffraction at a circular aperture.

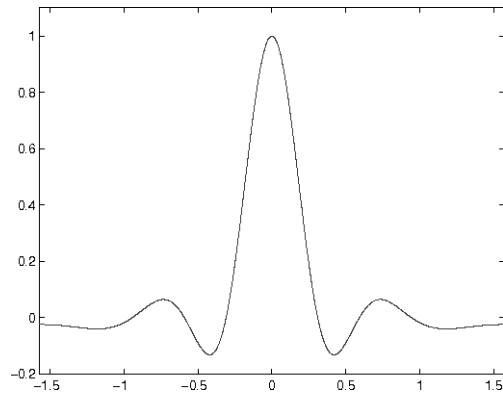


Fig. 4.7: Function (4.5) with $\frac{d}{\lambda} = 4$.

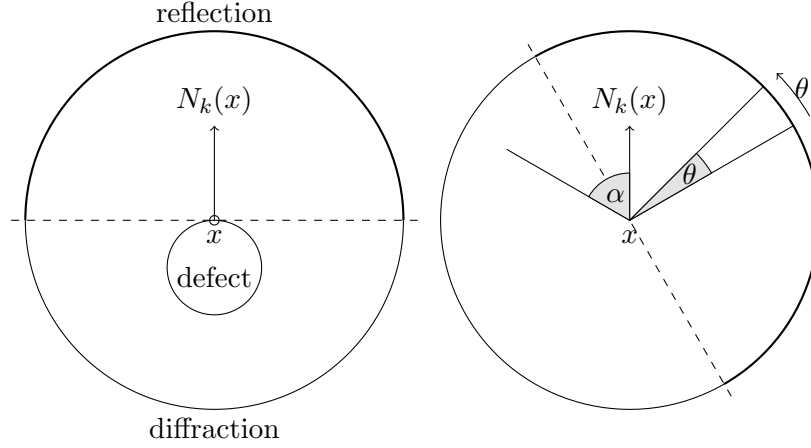


Fig. 4.8: Incident directions for reflection and diffraction (left) and behaviour of reflected amplitude depending on θ , amplitude is 0 in the thin half circle (right).

Note that $\Re\mathfrak{D}$ is usually small for reflecting points. This accords to the observation that in the case of a ToFD-measurement (see Subsection 2.2.3) the measured amplitudes will most likely belong to a diffraction of a defect.

After the ultrasound is reflected/diffracted by the defect, its amplitude will again decrease by the effects of attenuation and finally arrive at the receiver where it is measured. Thus, we can formulate our second model assumption which specifies the amplitude value by multiplying all considered effects.

Model Assumption 4.1.4:

Let all boundaries of D_k , $k = 1, \dots, K$ be regular. Then the amplitude of an ultrasound wave reflected or diffracted at

$$x \in E_{h,t} \cap \left(\bigcup_{k=1}^K \partial D_k \right)$$

that arrives at the receiver is given by

$$\mathbf{a}(s) \frac{e^{-\gamma \|x_E(h) - x\|_2 / c}}{\|x_E(h) - x\|_2} \Re\mathfrak{D}(\alpha) \frac{e^{-\gamma \|x_R(h) - x\|_2 / c}}{\|x_R(h) - x\|_2},$$

where the angle α can be derived by (4.4), γ is a material constant and s is given by

$$s = \frac{\|x - x_E(h)\|_2}{\|x - x_E(h)\|_2}.$$

Since $x \in E_{h,t}$ and Definition 4.1.1 this is equivalent to

$$e^{-\gamma t} \mathcal{A}(h, x) \quad \text{with} \quad \mathcal{A}(h, x) := \frac{\mathbf{a}(s) \Re\mathfrak{D}(\alpha)}{\|x_E(h) - x\|_2 \|x_R(h) - x\|_2}. \quad (4.7)$$

With the help of the model assumptions 4.1.2 and 4.1.4 we are now able to answer the two questions stated at the beginning of this subsection. To complete the presentation of our continuous model we have to combine all amplitudes that arrive at the same time t by summing them up. Assuming \mathcal{A} is integrable, we have to calculate the integral

$$e^{-\gamma t} \int_{E_{h,t} \cap (\cup_{k=1}^K \partial D_k)} \mathcal{A}(h, x) dx. \quad (4.8)$$

Here another problem arises as the manifold

$$E_{h,t} \cap (\cup_{k=1}^K \partial D_k) \quad (4.9)$$

might have different dimensions depending on h and t . Thus, we do not know if the integral should be understood as e.g. a line- or a surface-integral. In the discrete case of our model this will be no problem at all because the integral will be simplified to a sum over a discrete set of points but for the continuous case we have to make the assumption that the manifold in (4.9) does not form a surface in \mathbb{R}^3 . Then we can interpret the integral in (4.8) as a line-integral. This assumption fits into the consideration that the intersection between a defect and the ellipsoid $E_{h,t}$ will most likely form a line in \mathbb{R}^3 . The integral in (4.8) will be 0 if (4.9) is only a 0-dimensional manifold, e.g. a point. This means that the intersection is too small to cause any reflections or diffractions. Note that (4.9) is regular as all used sets are assumed to be regular. We can now formulate our first model.

Model 4.1.5 (The continuous geometrical model):

Let the boundaries of D_k , $k = 1, \dots, K$ be regular. Furthermore let

$$E_{h,t} \cap (\cup_{k=1}^K \partial D_k)$$

be a manifold of dimension 1 or 0 for all $h \in [0, 1]$, $t > 0$ and let \mathcal{A} be integrable over these manifolds in a line-integral sense. Then the continuous geometrical model function $\tilde{\mathbf{u}}$ is defined by

$$\tilde{\mathbf{u}}(h, t) := e^{-\gamma t} \int_{E_{h,t} \cap (\cup_{k=1}^K \partial D_k)} \mathcal{A}(h, x) dx,$$

where the integral is a line-integral. Since all D_k are disjoint, this can be rewritten as

$$\tilde{\mathbf{u}}(h, t) = e^{-\gamma t} \sum_{k=1}^K \int_{E_{h,t} \cap \partial D_k} \mathcal{A}(h, x) dx. \quad (4.10)$$

Note that model assumption 4.1.2 holds for (4.10) because all integrals will be 0 in the case that the manifold in (4.9) is the empty set. Moreover, the model includes

constructive or destructive interference by integration over $\mathcal{A}(h, x)$ where destructive interference means a change in the sign of $\mathcal{A}(h, x)$ on the region of integration. However, this model does only simulate the measured amplitudes and does not include the wave structure of ultrasound. In order to incorporate this aspect, we expand this model in the next subsection via a convolution.

4.1.2 Wave Convolution

The geometrical model can tell us at which time t an ultrasound wave is measured at the receiver and what its amplitude is. Let us first consider the simple example with $\tilde{\mathbf{u}}(h', t') = 1$ for fixed h', t' and $\tilde{\mathbf{u}}(h, t) = 0$ otherwise. This means, there is an incoming ultrasound wave for positions $x_E(h'), x_R(h')$ and time t' with amplitude 1. Now let the duration of this incoming impulse be τ , then the wave will be recorded for all $t \in [t', t' + \tau]$ while $\tilde{\mathbf{u}}(h, t) \neq 0$ only for $h = h'$ and $t = t'$. Thus, due to the missing wave structure, the geometrical model cannot simulate real data. To overcome this issue, we want to derive the model function \mathbf{u} that will also take the wave structure into account. Therefore we consider a function

$$f(t) : \mathbb{R}_+ \rightarrow \mathbb{R} \tag{4.11}$$

where in this work we define $\mathbb{R}_+ = [0, \infty)$ explicitly including 0. This function simulates the ultrasound impulse that is send out by the emitter at time $t = 0$. Note that one can expand f on \mathbb{R} by setting $f(t) = 0$ for $t < 0$. We will use this expansion implicitly in some of the following formulas. Remembering that we only consider homogeneous material we can assume that f is independent of h and that the incoming ultrasound waves that were measured at the receiver do not differ too much from f . Thus, considering the example from above with

$$\tilde{\mathbf{u}}(h, t) = \begin{cases} 1 & h = h', t = t', \\ 0 & \text{otherwise} \end{cases}$$

the model function \mathbf{u} should consist of a single impulse f that starts at time $t = t'$ for $h = h'$ with amplitude 1, i.e.

$$\mathbf{u}(h, t) = \begin{cases} \tilde{\mathbf{u}}(h', t')f(t - t') & h = h', \\ 0 & \text{otherwise.} \end{cases}$$

As above, we assume that the incoming wave has a duration of τ i.e. $\text{supp } f = [0, \tau]$. We obtain $\text{supp } \mathbf{u} = h' \times [t', t' + \tau]$ as desired. Now let us consider the general case. To continue the above considerations, whenever $\tilde{\mathbf{u}}(h, t) \neq 0$ holds, one should observe an

ultrasound impulse f that starts at time t with amplitude $\tilde{\mathbf{u}}(h, t)$. This can be done by convolution in the following way

$$\begin{aligned} \mathbf{u}(h, t) &= \int_0^{\infty} \tilde{\mathbf{u}}(h, \tau) f(t - \tau) d\tau \\ &= (\tilde{\mathbf{u}}(h, \cdot) * f)(t) \end{aligned}$$

which directly leads us to the continuous model.

Model 4.1.6 (The continuous model):

Using the assumptions and definitions of the continuous geometrical model 4.1.5 we define our continuous model function \mathbf{u} as

$$\mathbf{u}(h, t) = (\tilde{\mathbf{u}}(h, \cdot) * f)(t). \quad (4.12)$$

with a given ultrasound impulse function f . We assume that the convolution is well defined, what holds for e.g. $f, \tilde{\mathbf{u}}(h, \cdot) \in L_1(\mathbb{R})$.

We want to complete this section by giving two examples for the wave impulse f that we will mainly use in this work. The first one is a real valued Gabor function introduced in [74]

$$f(t) = e^{-\rho t^2} \cos(2\pi\psi t + \phi) \quad (4.13)$$

with a bandwidth factor $\rho > 0$, a center frequency ψ and a shift ϕ . The second example we want to consider is a raised cosine introduced in [7]

$$f_N(t) = \begin{cases} \left(1 + \cos \frac{2\pi\psi}{N} t\right) \cos(2\pi\psi t) & -N \frac{2}{\psi} \leq t \leq N \frac{2}{\psi}, \\ 0 & \text{otherwise} \end{cases} \quad (4.14)$$

with a center frequency ψ and a bandwidth factor N . Both functions (4.13) and (4.14) do not satisfy $f(t) = 0$ for $t < 0$ and thus they have to be shifted. The Gabor function (4.13) also has an infinite support and hence will be restricted to a finite interval in numerical applications. For a better representation we have neglected this in the formulas (4.14) and (4.13). Furthermore, we will only consider wave impulses f that are normalized due to their energy. This means the function will be multiplied by a constant such that $\|f\|_{L_2} = 1$ holds. Figure 4.9 shows two examples for the wave function f , a Gabor wave with $\rho = 0.75$, $\psi = 4$ and $\phi = 0$ (left) and as raised cosine with $\psi = 4$ and $N = 4$ (right).

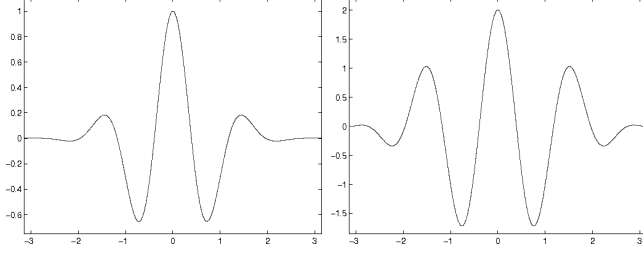


Fig. 4.9: Wave function f as Gabor function (left) and raised cosine (right).

4.2 Discretization

In the last section we have derived two continuous models, one using the geometric approach only and the second also including wave structures. Later in this chapter we will use these continuous formulas to show a connection between our models and the theory of inverse scattering. For numerical application of these models for e.g. the simulation of data we need a suitable discretization that will be developed in this section. Let us therefore introduce the following definition.

Definition 4.2.1:

For a set $\partial D \subset \mathbb{R}^n$ we call $\{P_l \in \partial D \mid l = 1, \dots, L\}$ an ε -discretization, if for all $x \in \partial D$ there exists an index l such that

$$\text{dist}_{\partial D}(x, P_l) := \min_{\mathbf{g} \in C^1} \left(\int_0^1 \|\mathbf{g}'(t)\|_2 dt \mid \mathbf{g}(t) : [0, 1] \rightarrow \partial D, \mathbf{g}(0) = x, \mathbf{g}(1) = P_l \right) \leq \varepsilon$$

holds with $\mathbf{g} \in C^1$, meaning that the shortest path on ∂D between x and P_l is at most of length ε . Define the set $C(P_l)$ as

$$C(P_l) := \{x \in \partial D \mid \text{dist}_{\partial D}(x, P_l) \leq \text{dist}_{\partial D}(x, P_{l'}) \text{ for all } l' \leq L\}.$$

Then $\cup_{l=1}^L C(P_l) = \partial D$, $C(P_l) \cap C(P_{l'})$ is a null set for $l \neq l'$ and for $x, y \in C(P_l)$ we obtain

$$\text{dist}_{\partial D}(x, y) \leq \text{dist}_{\partial D}(x, P_l) + \text{dist}_{\partial D}(y, P_l) \leq 2\varepsilon.$$

Thus the sets $C(P_l)$ with $l = 1, \dots, L$ form an ε -net of ∂D .

Given the ε -discretizations $\{P_l^k\}_{l=1}^{L_k}$ of the defect boundaries ∂D_k we approximate our continuous model by forming a Riemann-like sum of the integrals with the given sampling points. We define

$$t_{l,k,h} := \frac{\|P_l^k - x_E(h)\|_2 + \|P_l^k - x_R(h)\|_2}{c}$$

and, by Definition 4.1.1, note that $E_{h,t} \cap \{P_l^k\}_{l,k=1}^{L_k,K} \neq \emptyset$ only holds if $t = t_{l,k,h}$ for any l, k . Thus we obtain the set of sampling points $T = T_h = \text{unique}\{t_{l,k,h}\}_{k,l=1}^{K,L_k}$ for the convolution integral (4.12), where $\text{unique}\{\dots\}$ removes all but one element for values that appear multiple times. Then the inner integral

$$\int_{E_{h,t} \cap \partial D_k} \mathcal{A}(h, x) dx$$

in (4.10) is for $t_{l,k,h}$ sampled at the points P_l^k corresponding to $t_{l,k,h} \in T_h$. Because P_l^k is the representative for the set $C(P_l^k)$ which has a diameter of at most 2ε , we assume that the surface element $C(P_l^k)$ can be approximated by $\pi\varepsilon^2$. Altogether we obtain

Model 4.2.2 (The discrete model):

Given the ε -discretizations $\{P_l^k\}_{l=1}^{L_k}$ of the sets ∂D_k and the outer-pointing normals $\{N_l^k\}_{l=1}^{L_k}$ with $N_l^k = N_k(P_l^k)$, we then assume with the definitions from above that

$$u(h, t) \approx \pi\varepsilon^2 \sum_{\tau \in T_h} e^{-\gamma\tau} \sum_{k=1}^K \sum_{\substack{P_l^k \text{ with} \\ t_{l,k,h} = \tau}} \mathcal{A}(h, P_l^k, N_l^k) f(t - \tau) \quad (4.15)$$

$$= \pi\varepsilon^2 \sum_{k=1}^K \sum_{l=1}^{L_k} e^{-\gamma t_{l,k,h}} \mathcal{A}(h, P_l^k, N_l^k) f(t - t_{l,k,h}) \quad (4.16)$$

where we explicitly note the dependence of \mathcal{A} on N_l^k . As P_l^k is the representative of an area with a diameter of at most 2ε we remember our experiment with the circular aperture (see Figure 4.6) and set the diameter of the circular aperture $d = 2\varepsilon$ in (4.5).

The formula (4.15) will be useful for our considerations of the reconstruction method presented in this work as it separates the points in a nice manner while (4.16) can be used for implementations.

4.3 Model Validation

With the considerations of the last section we have derived a continuous model and a discrete implementation formula. Our model combines physical aspects in a heuristic way. However, we do not know anything about its approximation properties. Therefore, we want to check the quality of our model in this section. We will validate the model in two different manners. In the first subsection, we will show a connection to other approaches in (inverse) scattering, a field of research that has a widely accepted theory based on physical laws to model ultrasonic scattering problems among others. The second subsection will compare the output data of our model with given real data to expose similarities and differences.

4.3.1 Comparison to an Inverse Scattering Approach

The following considerations are based on Section 3.2 and the results in [39]. Note that we need the fundamental solutions of the Helmholtz equation in this subsection, thus all formulas have to be understood in the distributional sense. We first consider an ultrasound impulse $f \in L_2(\mathbb{R})$ that is inducted in an infinite homogeneous medium at time t_0 and at position x_0 . The generated wave \mathbf{u} has to satisfy the wave equation, meaning that

$$\frac{1}{c^2} \mathbf{u}_{tt}(x, t) - \Delta \mathbf{u}(x, t) = \delta(x - x_0) f(t - t_0),$$

where δ is the delta distribution. With the Fourier transform in time domain and $u(x, \omega) = \widehat{\mathbf{u}(x, t)}(x, \omega)$ we obtain

$$\frac{\omega^2}{c^2} u(x, \omega) - \Delta u(x, \omega) = \delta(x - x_0) e^{it_0 \omega} \hat{f}(\omega).$$

We note that this is the Helmholtz equation and for fixed ω the function u is given by the fundamental solution. Thus

$$u(x, \omega) = e^{it_0 \omega} \hat{f}(\omega) \frac{e^{i\frac{\omega}{c} \|x - x_0\|_2}}{4\pi \|x - x_0\|_2}$$

and with the inverse Fourier transform,

$$\begin{aligned} \mathbf{u}(x, t) &= \frac{1}{4\pi \|x - x_0\|_2} f(t - t_0) * \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega(\frac{\|x - x_0\|_2}{c} - t)} d\omega \right) \\ &= \frac{1}{4\pi \|x - x_0\|_2} f(t - t_0) * \delta\left(t - \frac{\|x - x_0\|_2}{c}\right) \\ &= \frac{1}{4\pi \|x - x_0\|_2} f\left(t - t_0 - \frac{\|x - x_0\|_2}{c}\right). \end{aligned} \quad (4.17)$$

Let us now consider an emitter at position $x_E(h)$ for fixed $h \in [0, 1]$ that emits an ultrasound impulse f at time $t_0 = 0$ and the defects D_1, \dots, D_K . Then the generated ultrasound wave \mathbf{u} is given by the sum of the incident wave \mathbf{u}^i and the scattered wave \mathbf{u}^s . Note that the incident wave is directly given by formula (4.17) as

$$\mathbf{u}^i(x, t) = \frac{1}{4\pi \|x - x_E(h)\|_2} f\left(t - \frac{\|x - x_E(h)\|_2}{c}\right).$$

However, the scattered wave depends non-linearly on itself, meaning that $\mathbf{u}^s = \mathbf{u}^s(\mathbf{u}^i, \mathbf{u}^s)$ because of multi-scattering effects. To approximate the scattered wave \mathbf{u}^s we use the so called *Born approximation* that neglects those effects and considers $\mathbf{u}^s \approx \mathbf{u}^s(\mathbf{u}^i)$ [7, 39]. For an error bound of the Born approximation we refer to [75]. Given that \mathbf{u}^s only

depends on the incident wave \mathbf{u}^i we can consider every point on ∂D_k as a single *secondary* ultrasound source that is stimulated by \mathbf{u}^i , so its generated wave is given by (4.17) with $f = \mathbf{u}^i(x, \cdot)$. Then the scattered wave is the superposition of all secondary sources. We obtain

$$\begin{aligned} \mathbf{u}^s(x, t) &= \sum_{k=1}^K \int_{\partial D_k} \frac{1}{4\pi \|x - y\|_2} \mathbf{u}^i \left(y, t - \frac{\|x - y\|_2}{c} \right) dy \\ &= \sum_{k=1}^K \int_{\partial D_k} \frac{1}{16\pi^2 \|x - y\|_2 \|y - x_E(h)\|_2} f \left(t - \frac{\|x - y\|_2 + \|y - x_E(h)\|_2}{c} \right) dy. \end{aligned}$$

Thus, the data that is recorded by the receiver at position $x_R(h)$ can be simulated by

$$\begin{aligned} \mathbf{u}(x_R(h), t) &= \mathbf{u}^i(x_R(h), t) + \mathbf{u}^s(x_R(h), t) \\ &= \frac{1}{4\pi \|x_R(h) - x_E(h)\|_2} f \left(t - \frac{\|x_R(h) - x_E(h)\|_2}{c} \right) \\ &\quad + \sum_{k=1}^K \int_{\partial D_k} \frac{f \left(t - \frac{\|x_R(h) - y\|_2 + \|y - x_E(h)\|_2}{c} \right)}{16\pi^2 \|x_R(h) - y\|_2 \|y - x_E(h)\|_2} dy, \end{aligned}$$

where the first term \mathbf{u}^i simulates the lateral wave and thus will be neglected in the following considerations. We now assume that there exists a regular parametrisation $\Gamma : [t_0, t_1] \times [0, 1] \rightarrow \partial D$ with $\Gamma(t, \cdot) : [0, 1] \rightarrow \partial D \cap E_{h,t}$. It follows that $\partial D \cap E_{h,t} = \emptyset$ for $t > t_1$ or $t < t_0$. Let Γ_1, Γ_2 be the partial derivatives of $\Gamma(t, s)$ with respect to s and t , and let $\theta(t, s) = \sphericalangle(\Gamma_1(t, s), \Gamma_2(t, s))$ be the angle between the two derivatives at the point (t, s) , then we obtain

$$\begin{aligned} \mathbf{u}(x_R(h), t) &= \sum_{k=1}^K \int_{t_0}^{t_1} \int_0^1 \frac{\|(\Gamma_1 \times \Gamma_2)(\tau, s)\|_2 f \left(t - \frac{\|x_R(h) - \Gamma(\tau, s)\|_2 + \|\Gamma(\tau, s) - x_E(h)\|_2}{c} \right)}{16\pi^2 \|x_R(h) - \Gamma(\tau, s)\|_2 \|\Gamma(\tau, s) - x_E(h)\|_2} ds d\tau \\ &= \sum_{k=1}^K \int_{t_0}^{t_1} \int_0^1 \frac{\|\Gamma_1(\tau, s)\|_2 \|\Gamma_2(\tau, s)\|_2 |\sin(\theta(\tau, s))|}{16\pi^2 \|x_R(h) - \Gamma(\tau, s)\|_2 \|\Gamma(\tau, s) - x_E(h)\|_2} \\ &\quad \cdot f \left(t - \frac{\|x_R(h) - \Gamma(\tau, s)\|_2 + \|\Gamma(\tau, s) - x_E(h)\|_2}{c} \right) ds d\tau \\ &= \sum_{k=1}^K \int_{t_0}^{t_1} \int_{\partial D_k \cap E_{h,t}} \frac{\|\Gamma_1(\Gamma^{-1}(y))\|_2 |\sin(\theta(\Gamma^{-1}(y)))| f(t - \tau)}{16\pi^2 \|x_R(h) - y\|_2 \|y - x_E(h)\|_2} dy d\tau \\ &= f * \sum_{k=1}^K \int_{\partial D_k \cap E_{h,t}} \frac{\|\Gamma_1(\Gamma^{-1}(y))\|_2 |\sin(\theta(\Gamma^{-1}(y)))|}{16\pi^2 \|x_R(h) - y\|_2 \|y - x_E(h)\|_2} dy. \end{aligned}$$

We observe that the difference between our new model and the approach via inverse scattering using the Born approximation especially regards the amplitude function. A comparison with model 4.1.5 yields

$$\begin{aligned} \text{new model} & \leftrightarrow \text{scattering with Born approximation,} \\ \frac{e^{-\gamma t} \mathbf{a}(s) \Re \mathcal{D}(\alpha)}{\|x_R(h) - x\|_2 \|x - x_E(h)\|_2} & \leftrightarrow \frac{\|\Gamma_1(\Gamma^{-1}(x))\|_2 |\sin(\theta(\Gamma^{-1}(x)))|}{16\pi^2 \|x_R(h) - x\|_2 \|x - x_E(h)\|_2}. \end{aligned}$$

We note that, because Γ is regular, $\|\Gamma_1(\Gamma^{-1}(x))\|_2 |\sin(\theta(\Gamma^{-1}(x)))| > 0$ holds. Hence the Born approximation does neither consider a phase shift (e.g. by reflection) nor destructive interference of the amplitudes. As a side note we also want to mention that the inversion method introduced in this work does not directly use the structure of the amplitude function, so it can be applied to our new model as well as to the model obtained by the scattering approach via Born approximation.

4.3.2 Numerical Validation

In the last subsection we have shown the connection between our model and an inverse scattering approach that can be seen as some kind of *theoretical* validation. Next we are going to compare our model output data with real data, hence perform a *practical* validation. Besides the real output data, we need some technical data on the process of data requisition. For most output data it is no problem at all to get the emitter type, the frequency of the impulse signal and other testing settings. However, a priori information about the defects in the material are not available. As the defect is inside a steel tube, it is often not accessible and one needs to disassemble the component. This may be impossible as the tube might e.g. be already in use. Even if one is able to disassemble it this process might destroy the defect. In order to overcome this problem, special types of defects are produced to obtain the corresponding real data. In this way we can evaluate if our model data and the pattern shown by real data fit together. Thereby we will consider three types of defects, firstly a pore defect with a ToFD set-up, secondly a crack also examined with a ToFD set-up and at last a half-spheroid back wall defect in a wall thickness measurement.

Let us start with the real data from ToFD measurements with pore defects shown in Figure 4.10. Besides the lateral wave and the back wall echo, which we will not take into account here, we observe parabolic like structures in the data (marked). These are typical reflection and diffraction signals of pore defects in the material as emitter and receiver first move closer to the defect, meaning the signals are measured earlier in time and more stronger reaching its tip at the apex of the parabolic structure and then begin to decrease as the probes move away from the pore. Note that the data also contains smaller structures and noise that are mostly caused by scattering on minor structures. We will consider this noise later in this section.

To reproduce the output data for a pore by our model, we have simulated a spherical defect with a diameter of 1.5mm that is 5mm under the material surface. Equivalently to the real ToFD set-up, emitter and receiver have been placed on lines at the left and the right of the defect moving along these lines for a distance of 44cm where they pass the nearest point to the defect after 22cm. a-scans have been recorded every 1mm. The function \mathbf{a} is simulating a focussing emitter such that the pore is slightly out of focus and as an impulse function we have taken the Gabor wave (see 4.13) with bandwidth factor 0.75MHz^2 , a center frequency of 1Mhz and no shift. For the time discretization we have chosen a step size of $0.05\mu\text{s}$ and the ε -discretization has been computed for $\varepsilon = 0.1$. The obtained output data can be seen in Figure 4.11. We note that the model produces the same parabolic like structures as observed in the real data, thus seems to be a good approximation for pore defects.

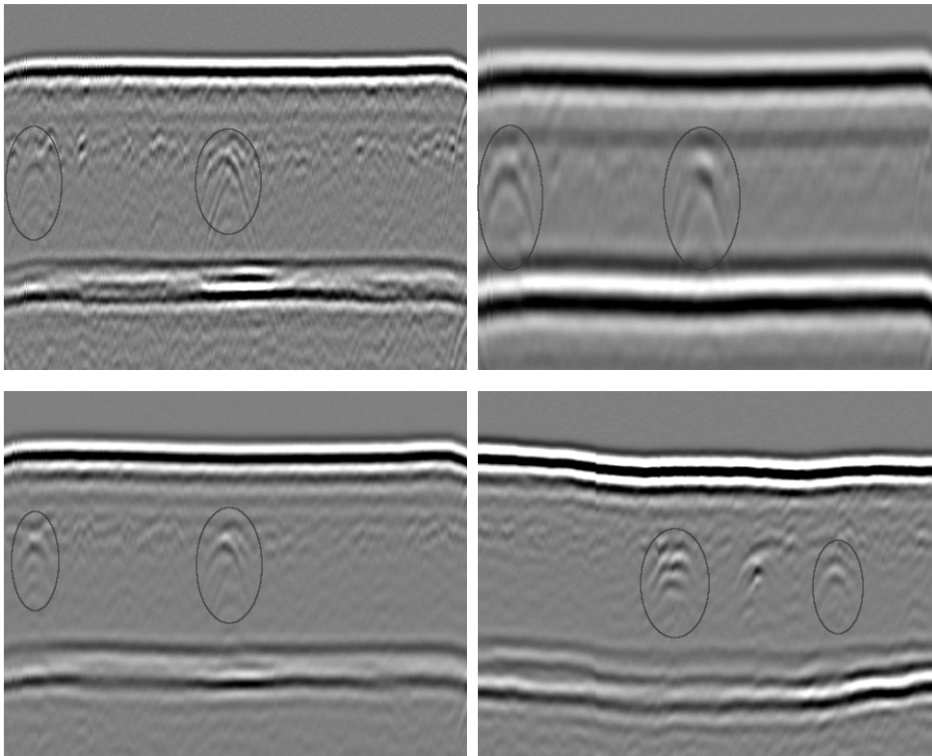


Fig. 4.10: Different real output data of pore defects in ToFD.

Now let us have a look at the output data produced by crack defects in ToFD measurements. Real output data is given in Figure 4.12. We observe that the measured signals are mostly concentrated along a line that seems to follow the shape of the crack. Figure 4.13 shows the simulated output data for a ToFD measurement with the same parameters as above but an impulse Gabor function with bandwidth factor 1.25 and frequency 2.75. Now we have placed an 8mm long horizontal crack at the end of the component

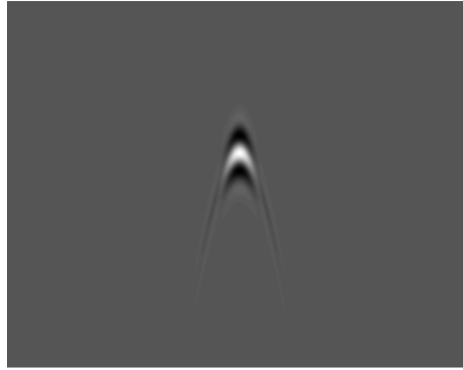


Fig. 4.11: Modelled ToFD output data for a pore defect.

right below the surface. Similarly to real data, we can observe the same concentration of signals along a horizontal line.

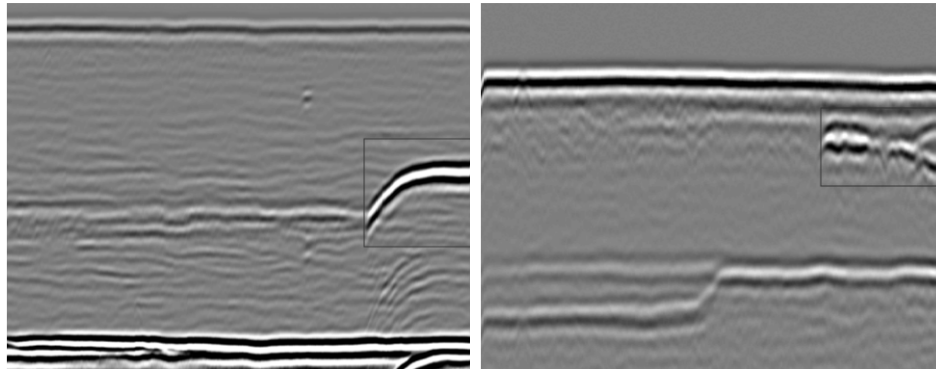


Fig. 4.12: Different real output data of crack defects in ToFD.

After having compared the output data of our model with two set-ups for the ToFD method, we now consider a back wall thickness measurement. Therefore Figure 4.14 shows real output data for different damaged back walls. Again, the data also contains a surface reflection and multiple reflections of the back wall which will not be taken into account. We observe that a defect of the tube produces a breakup of the back wall echo in the output data. Depending on the size and form of the defect this breakup can be abrupt or slowly fading. Instead of the back wall signal one may receive a defect reflection also appearing abruptly or slowly fading.

To simulate the output data of a back wall thickness measurement we have modelled a probe, being emitter and receiver in once, that moves over the surface of a 5cm thick component along a 30cm long line taking an a-scan every 1mm . The probe focus is directly targeting the back wall and the impulse is simulated with a Gabor function with bandwidth factor 1.25, frequency 2.75 and no shift. The discretization step size in time is $0.05\mu\text{s}$ and we choose $\varepsilon = 0.1$. A smooth thin but 16cm long and 5mm high



Fig. 4.13: Modelled ToFD output data for a crack defect.

defect is placed in the back wall. The resulting data is shown in Figure 4.15. Here one can observe the same effects as in real data. The back wall echo vanishes and a signal reflection of the defect appears. Note that due to the smoothness of the defect the signals fade out/in very smoothly.

As we have seen, our model is able to simulate the same effects that appear in real data. But besides signals caused by reflection and diffraction at defect boundaries, real data also contains a significant rate of noise caused by e.g. multiple scattering or material inhomogeneities. This noise can cause problems in the inversion process as ill-posed problems typically are very prone to it. Thus, to counteract this problem, we have to understand the structure of this noise. One can observe by the different real data examples (see Figure 4.10, 4.12 and 4.14) that the noise has also a wave like structure. This is consistent with the above consideration that the noise is caused by multiple scattering and inhomogeneities. Thus, noise seems to be randomly incoming ultrasound waves with small amplitude. We assume that our model can be expanded with such noise by adding Gaussian noise to the geometrical model function \tilde{u} . Figure 4.16 shows the modelled data with added Gaussian noise (variance 0.01). Indeed, we observe that the modelled noise produces effects similar to the noise on real data.

After the validation of our new model we will turn our attention to the inverse problem in the next chapters. We have seen similarities of our model to the field of inverse scattering as well as to real output data. Thus it can serve as a fundament for our inversion method that we will present in this work. Note that there will be further and more detailed numerical examples in Chapter 9 where we will numerically analyse our reconstruction method.

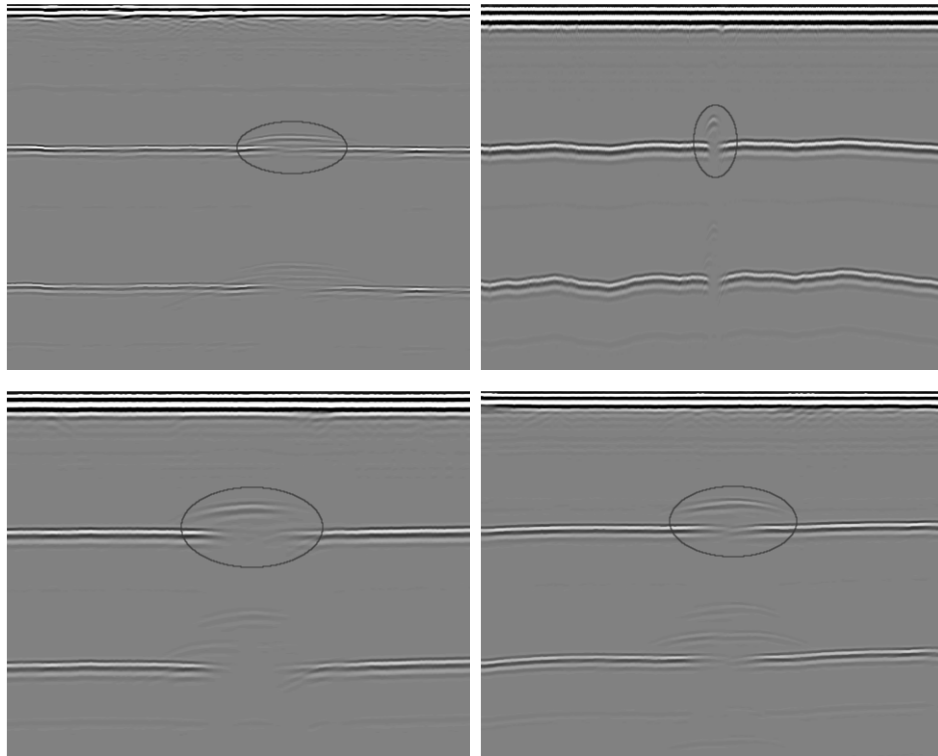


Fig. 4.14: Different real output data of back wall thickness measurement.



Fig. 4.15: Modelled back wall thickness measurement.

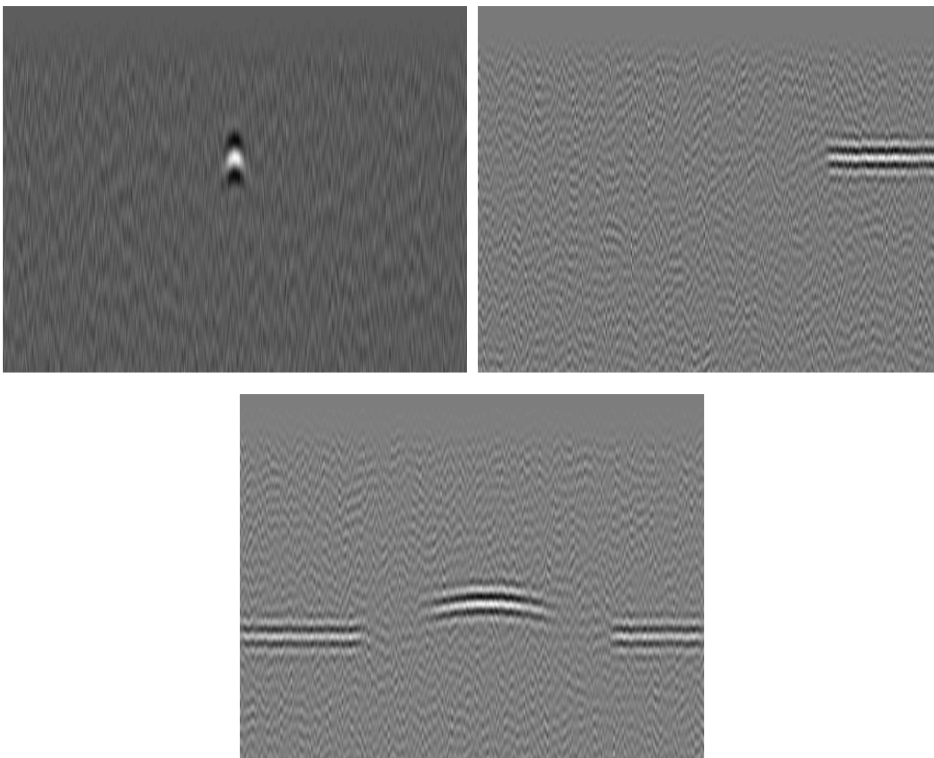


Fig. 4.16: Modelled data with Gaussian noise on \tilde{u} .

5 Inversion - an Overview

After we have considered several topics about the direct problem in the previous chapters, we are now prepared to introduce our new inversion method. In Chapter 1 the problem and the necessity for a reliable algorithm was introduced. We learned about the physical background and non-destructive testing techniques in Chapter 2. The actually used methods in non-destructive testing as well as in nearby fields of research were presented in Chapter 3. Thereby we also commented on the difficulties and weaknesses of some methods and why they cannot be adapted easily to our problem. Finally, in Chapter 4 we designed a new model based on heuristics that will be the foundation for our new inversion method.

Our algorithm can be divided into three separated steps. To give the reader an idea what the aim of each inversion step is and how they are combined to a complete reconstruction method, this chapter will give a brief overview of the complete algorithm and introduce the considerations behind those steps. Each step will later be discussed in detail in Chapters 6 to 8.

To introduce our inverse method we consider a model function \mathbf{u} of the form

$$\mathbf{u}(h, t) = \left(f(t) * \left(\sum_{k=1}^K \int_{\partial D_k \cap E_{h,t}} a(t, y, h) dy + N(t, h) \right) \right) (t) \quad (5.1)$$

with the notations of the last chapters, an amplitude function a and Gaussian noise $N(t, h)$. Note that our model with $a = e^{-\gamma t} \mathcal{A}(y, h)$ as well as the inverse scattering model using Born approximation can be described by 5.1. Our reconstruction algorithm can be applied to most problems that can be modelled by (5.1). Indeed the only assumption we need, is that

$$\left| \sum_{k=1}^K \int_{\partial D_k \cap E_{h,t}} a(t, y, h) dy + N(t, h) \right|$$

generally is big in relation to the noise level if $\partial D_k \cap E_{h,t} \neq \emptyset$ for at least one k . This means, if there is a reflection or diffraction caused by the defect, there is a considerable increase of the amplitude. If $\partial D_k \cap E_{h,t} = \emptyset$, the amplitude then only consists of the noise $N(t, h)$ and should stay low. This assumption assures that the important signal structures can be separated from the noise, otherwise we would not be able to reconstruct

all information transferred by this amplitude function. Using this assumption we have to solve the following inverse problem, namely

The Inverse Problem 5.1.1:

Given the function \mathbf{u} or samples of it defined by (5.1), reconstruct $\{\partial D_k\}_k^K$ where $K \in \mathbb{N}$ is also unknown.

Let us now have a closer look at formula (5.1). To reconstruct ∂D_k out of (5.1) we first need to invert the convolution, secondly the summation and at last the integral. We use separate methods for each of these steps. The combination of these methods will present the complete reconstruction algorithm.

We will start considering the deconvolution of formula (5.1). Therefore, we assume that there are only few and small defects in the material, which is a realistic assumption because today's manufacturing is quite good. As a consequence, we note that few defects can only cause few significant amplitudes in the data. For most h and t the intersection $\partial D_k \cap E_{h,t}$ will be empty and thus the amplitude function will only consist of noise. Hence we can rewrite \mathbf{u} as

$$\mathbf{u} = f * (X + N) \tag{5.2}$$

where N is noise and X is sparse. The inverse problem to reconstruct X now becomes a sparse deconvolution problem and assuming that $f * N$ is negligible we are searching for a X' such that $\mathbf{u} \approx f * X'$. Such a problem can be solved by using greedy algorithms like Matching Pursuit (MP) or Orthogonal Matching Pursuit (OMP) [76, 77] but one may also consider the Prony method [78, 79]. The details of all methods and their adaptation to our inverse problem will be presented in Chapter 6 that is based on our publication [80]. These techniques try to reconstruct a sparse solution X' of problem (5.2). We will later see that for the practical problem of defect reconstruction the most important information of X' is given by its support and not by its values which are often strongly influenced by noise. Thus we consider the essential support of the reconstruction X' that is related to the support of significant values of $X + N$. Due to the assumption that the amplitude of the noise N is small compared to significant structures of X this support will be also related to the essential support of X . In this way, we may assume that $\text{supp } X' \approx \text{supp } X$ holds and use this information for the next inversion step. Note that the first step approximates the usually small essential support of X for given data \mathbf{u} . Hence we may understand this deconvolution step as denoising and information extraction of \mathbf{u} .

As a result of the first step of our inversion method we obtain the support of the sparse deconvolved and denoised data

$$\sum_{k=1}^K \int_{\partial D_k \cap E_{h,t}} a(t, y, h) dy. \tag{5.3}$$

Let us first assume $K = 2$, i.e. there are two defects in the material which should also have great distance to each other. Now it follows that most likely if $\partial D_1 \cap E_{h,t} \neq \emptyset$ then $\partial D_2 \cap E_{h,t} = \emptyset$ and otherwise, meaning that only one element of the sum in (5.3) is non-zero at a time. If one defect causes a signal at time t and position h then, because of the great distance between both defects, the other one will surely be too far away to cause a signal at the same time t . In this scenario the support of (5.3) will consist of two parts, where for the first part $\partial D_1 \cap E_{h,t} \neq \emptyset$ will hold while for the second part we obtain $\partial D_2 \cap E_{h,t} \neq \emptyset$. Moreover, because of the long distance between both defects, these parts can be separated forming two clusters of the support of (5.3). We can easily find those clusters using a clustering algorithm on the given support. Then each cluster contains a part of the signal that is caused by only one defect and can be used to apply our defect fitting algorithm as the last step of the inversion method. Let us now consider a more general case where the number of defects K is unknown. In this case one has to use a clustering algorithm where the number of clusters is not predetermined. The number of returned clusters will then correspond with the numbers of defects in the material. Using such algorithms it is possible to reconstruct K . Now, considering also the case of two defects near to each other, the clusters will surely intersect, what means they can no longer be separated by a standard clustering algorithm. We will present ideas to adapt those clustering algorithms using also geometrical information of the support, as the two defects may not be separated in the coordinate plane but the structure of the caused signals allows a reconstruction of both clusters. Thus, this step of the inversion aims to reconstruct the number of defects K and attaches each signal to its source defect. The details of these step will be presented in Chapter 7.

In the last step of our reconstruction method, we will reconstruct the defects by using the information obtained from the previous steps. As we have separated the defects in step 2 we can now use a reconstruction method that only considers one defect. We apply it to each cluster on its own. Thus, the problem can be stated as follows. For given

$$\int_{\partial D_k \cap E_{h,t}} a(t, y, h) dy \quad (5.4)$$

and fixed $k \leq K$, how can we reconstruct the set ∂D_k ? Note that the reconstructed amplitude is not reliable because of the noisy measurement. Furthermore, the function a depends on ∂D and thus the unknown defect appears both in the integral and in its domain making the inversion problem even harder. For these reasons, we will only use the information about the support of (5.4), i.e. the pairs (t, h) such that (5.4) is not equal 0. One may consider the information given by the amplitude in reconstruction problems where more data is given. It may also be possible to use the structure of a in a fourth step to improve the obtained solution. However, our algorithm does not make

use of the information provided in the amplitudes. Furthermore, instead of using the support of (5.4) we use a slightly relaxed information, namely that

$$\partial D \cap E_{h,t} \neq \emptyset \text{ for all } (h,t) \in \text{Cl}$$

where we omit the k in $\partial D = \partial D_k$ and Cl is the cluster derived by the second step of our reconstruction method. Again we use the assumption that the defect ∂D will most likely be small. This leads to the minimization problem

$$\min_D D \quad \text{with} \quad \partial D \cap E_{h,t} \neq \emptyset \quad \forall (h,t) \in \text{Cl} \quad (5.5)$$

where the minimization has to be defined in a reasonable way. Although the constraints are non-convex we will use their special structure to develop an algorithm that starts with a big star-shaped defect and reduces it iteratively until it stops in a local minimum of (5.5). The details of step 3 will be presented in Chapter 8.

Now we have presented a short overview of the methods and ideas used in our inversion method. A graphical illustration is shown in Figure 5.1. The next chapters will discuss all steps of our algorithm in details. We strongly recommend to keep the ideas of this chapter in mind or reread it whenever necessary as the aims of each step and their collaboration will help to understand the following considerations.

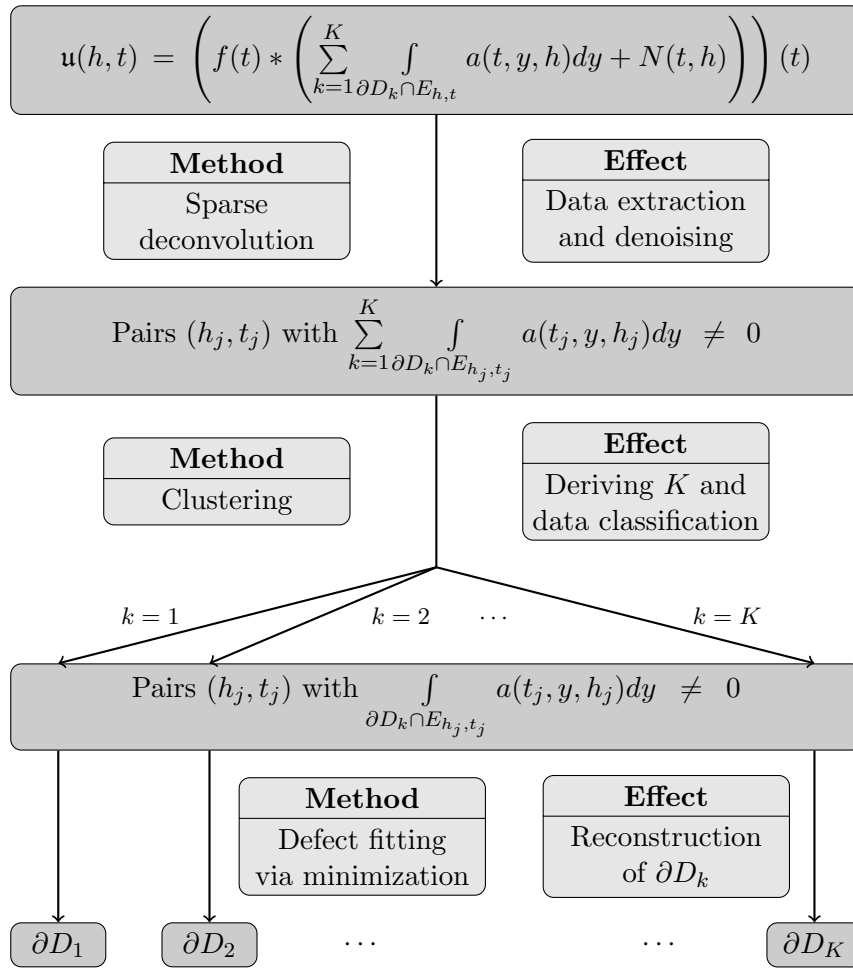


Fig. 5.1: Illustration of the algorithm.

6 Inversion 1 - Deconvolution

In this chapter, based on [80], we want to introduce the first step of our reconstruction technique. Although the ideas of this step have been presented in the last chapter, a short summary will be given for an easier understanding. Therefore consider our model

$$\mathbf{u}(h, t) = \left(f(t) * \left(\sum_{k=1}^K \int_{\partial D_k \cap E_{h,t}} a(t, y, h) dy + N(t, h) \right) \right) (t). \quad (6.1)$$

In order to reconstruct the defects in the material, our first aim is to deconvolve the data and remove the noise $N(t, h)$. We assume only few and small defects. This is a reasonable assumption due to today's quality that can be achieved in steel production. As a direct consequence we obtain that

$$\sum_{k=1}^K \int_{\partial D_k \cap E_{h,t}} a(t, y, h) dy \quad (6.2)$$

is a sparse signal while the noise $N(t, h)$ is not. This is a powerful constraint that we want to exploit by using sparse deconvolution methods. These methods are especially designed to find a sparse solution X of the problem

$$\mathbf{u} = f * X.$$

As the reflected and diffracted signals caused by the defects will most likely form the main components of the signal \mathbf{u} , the reconstructed solution will be a good approximation to (6.2). Moreover, the noise $N(t, h)$ is a non sparse signal with low amplitudes and thus it has only little effect on the solution and most of it will be filtered out. Hence, sparse deconvolution methods satisfy all requirements to our first reconstruction step as they provide a technique to deconvolve and simultaneously denoise the data. We will discuss those methods in the next section in detail.

Note that until now we have implicitly assumed that the wave function f in (6.1) is known to allow a deconvolution of the data. This is often not the case in practical applications. In the second section of this chapter we will therefore discuss techniques to solve a blind deconvolution problem where f is unknown. This can be done by iteratively solving the sparse deconvolution problem with an approximation $\tilde{f} \approx f$ and then applying an

update step to \tilde{f} . However, with the blind deconvolution problem further difficulties appear that have to be considered, as e.g. a starting guess for \tilde{f} is needed.

6.1 Sparse Deconvolution

Let us first consider a general sparse deconvolution problem, namely to reconstruct X from given data

$$u = f * X$$

where f is known and X is sparse. Furthermore we assume that at least an upper bound $L \in \mathbb{N}$ for the number of non-zero elements in X is given. Today there is a wide range of algorithms to solve such a sparse deconvolution problem as e.g. the Prony method [79, 80], the MUSIC method [81] or reconstruction via l_1 -minimization [82]. As each technique possesses its own advantages and disadvantages one has to choose a method depending on the requirements that occur with the problem. In case of ultrasonic non-destructive testing these are stability, even with highly noised data, and the need for a computational efficient implementation to allow a fast defect reconstruction. Both requirements reduce the number of possible algorithms drastically as the reconstruction of a sparse solution is complex and thus many algorithms have high cost and/or are prone to noise. However, there is one class of algorithms that is especially designed to (approximatively) solve problems with high complexity in an efficient way, the so called “Greedy”-methods [83, 84]. Roughly said, these methods try to find a *global* solution of a problem by solving multiple *local* problems and combining them. Considering sparse deconvolution this e.g. means that the algorithm tries to find an L -sparse solution by deriving L times a 1-sparse solutions of related problems. As a disadvantage that comes at hand, Greedy methods do not always recover the exact solution but often return a good approximation of it. Thus, the resolution of those algorithms might be less accurate than other methods. However, in ultrasonic non-destructive testing this is less important than a fast and stable implementation. Hence, the greedy approach is our method of choice. In this work we will make use of two greedy algorithms, namely the *Matching Pursuit* (MP) and the *orthogonal Matching Pursuit* (OMP) [76, 77]. We will introduce these algorithms in the following subsection.

6.1.1 (Orthogonal) Matching Pursuit

Two Greedy methods that are based on very simple ideas and therefore allow a fast implementation, are the Matching Pursuit (MP) and its extension the orthogonal Matching Pursuit (OMP). Both algorithms and modified versions of them have been considered earlier in ultrasonic non-destructive testing, see [85–87]. However, in [80] we applied them for the first time in connection with a blind deconvolution approach when the ultrasound impulse f is unknown. Furthermore, to the authors best knowledge, this work

is the first to combine Matching Pursuit with additional methods to design a complete reconstruction algorithm.

In this subsection we will first present the two techniques MP and OMP independently from our application. Later in this chapter, we will discuss their adaption to ultrasonic non-destructive testing. Since OMP is based on MP, we will start with the MP algorithm. We consider a function \mathbf{u} in a Hilbert space \mathcal{H} with an inner product $\langle \cdot, \cdot \rangle$ and the corresponding norm $\|\cdot\|_{\mathcal{H}}$. Let $\mathcal{D} = \{f_1, \dots, f_D\}$ be a dictionary with $f_k \in \mathcal{H}$ and $\|f_k\|_{\mathcal{H}} = 1$ for $k = 1, \dots, D$. MP now tries to find the best L -term approximation of \mathbf{u} according to the dictionary \mathcal{D} , i.e. $1 \leq j_1 < \dots < j_L \leq D$ and $x_1, \dots, x_L \in \mathbb{C}$ that minimize

$$\left\| \mathbf{u} - \sum_{k=1}^L x_k f_{j_k} \right\|_{\mathcal{H}}.$$

MP uses L iterations to recover a good L -term approximation. In each iteration it chooses *greedily* one element of the dictionary \mathcal{D} that matches best to the remaining residual. Thus, the algorithm works as follows:

1. Let a function $\mathbf{u} \in \mathcal{H}$, a dictionary \mathcal{D} and $L \in \mathbb{N}$ with $L \ll \#\mathcal{D}$ be given. Initialize $k := 1$, $r_1 := \mathbf{u}$ and iterate steps 2 and 3 until $k = L$.

2. Determine the element $f_{j_k} \in \mathcal{D}$ that correlates most with the residual r_k , i.e. solve

$$j_k := \arg \max_{j=1, \dots, D} |\langle r_k, f_j \rangle|.$$

3. Update $x_k := \langle r_k, f_{j_k} \rangle$, $r_{k+1} := r_k - x_k f_{j_k}$ and $k := k + 1$.

4. Return $\{j_k\}_{k=1}^L$ and $\{x_k\}_{k=1}^L$.

We know that

$$\begin{aligned} \min_j \min_a \|r_k - a f_j\|_{\mathcal{H}}^2 &= \min_j \|r_k - \langle r_k, f_j \rangle f_j\|_{\mathcal{H}}^2 \\ &= \min_j (\|r_k\|_{\mathcal{H}}^2 - |\langle r_k, f_j \rangle|^2) \\ &= \|r_k\|_{\mathcal{H}}^2 - \max_j |\langle r_k, f_j \rangle|^2, \end{aligned} \tag{6.3}$$

and thus the MP solves in each iteration the local problem (6.3) to approximate the solution of the global problem

$$\min_{\substack{x_k, j_k \\ k=1, \dots, L}} \left\| \mathbf{u} - \sum_{k=1}^L x_k f_{j_k} \right\|_{\mathcal{H}}. \tag{6.4}$$

However, the reconstruction quality of MP strongly depends on the correlation between the active dictionary elements, i.e. between the elements $\{f_{j_k}\}_{k=1}^L$. This can be seen easily in the following example. Consider the function

$$\mathbf{u} = \sum_{k=1}^L f_k.$$

Then the optimal solution of (6.4) is given by $j_k = k$ and $x_k = 1$ for $k = 1, \dots, L$. Assume the first dictionary element chosen by MP is $j_1 = 1$. Then the first coefficient reconstructed is given by

$$\langle \mathbf{u}, f_1 \rangle = 1 + \sum_{k=2}^L \langle f_1, f_k \rangle,$$

and this term is not necessarily 1 but depends on the correlation between the dictionary elements. It recovers the correct coefficients only if the dictionary elements are orthogonal. As an extension of MP, the orthogonal Matching Pursuit tries to overcome this problem. Compared to MP, there is an additional optimization of the parameters x_k in each iteration. In detail, OMP works as follows:

1. Let a function $\mathbf{u} \in \mathcal{H}$, a dictionary \mathcal{D} and $L \in \mathbb{N}$ with $L \ll \#\mathcal{D}$ be given. Initialize $k := 1$, $r_1 := \mathbf{u}$ and iterate step 2-4 until $k = L$.
2. Determine the element $f_{j_k} \in \mathcal{D}$ that correlates most with the residual r_k , i.e. solve

$$j_k := \arg \max_{j=1, \dots, D} |\langle r_k, f_j \rangle|.$$

3. Calculate $\{x_i\}_{i=1}^k$ by solving $\min_{x_i} \|\mathbf{u} - \sum_{i=1}^k x_i f_{j_i}\|_{\mathcal{H}}$.
4. Update $r_{k+1} := \mathbf{u} - \sum_{i=1}^k x_i f_{j_i}$ and $k := k + 1$.
5. Return $\{j_k\}_{k=1}^L$ and $\{x_k\}_{k=1}^L$.

Thus, OMP recomputes all coefficients in each iteration to take the correlations between active dictionary elements into account.

Besides OMP there are plenty of algorithms based on MP trying to improve accuracy, stability or runtime by adding additional ideas, see e.g. [86–90]. Here we will only consider the use of MP and OMP as the detailed discussion of all variations of MP would go beyond the scope of this work, but note that one could easily replace MP and OMP in our reconstruction method by other algorithms that approximately solve the sparse deconvolution problem.

6.1.2 MP and OMP for Sparse Deconvolution of NDT Data

Now we understand how MP and OMP work in a general framework. Let us analyse how it adapts to the special set-up of sparse deconvolution of ultrasonic NDT data. Therefore we consider an a-scan \mathbf{u} , i.e. ultrasound data received from only one measurement. Since the convolution of our model is only in the time dimension, all results obtained here can also be used to deconvolve data containing more a-scans by applying the methods on each a-scan separately. Now, assume that the ultrasound data

$$\mathbf{u} = f * X \quad (6.5)$$

is sampled on a grid $0 \leq t_1 < \dots < t_N < \infty$ and the impulse function f is known. Thus the function vector $(\mathbf{u}(t_k))_{k=1}^N$ is given where we suppress the dependence on h as the convolution is only performed in the parameter t . We first consider the noiseless case where the number of ultrasound impulses L in \mathbf{u} is known and assume that \mathbf{u} is of the form

$$\mathbf{u} = \sum_{k=1}^L x_k f(\cdot - \tau_k)$$

with shift parameters τ_1, \dots, τ_L and $x_k \in \mathbb{R}$ for $k = 1, \dots, L$. Thus the optimal dictionary for a sparse representation of \mathbf{u} would be $\mathcal{D} = \{f(\cdot - \tau_k)\}_{k=1}^L$. Unfortunately, the exact shifts are not known. For this reason we need to design a dictionary that does not necessarily include the correct shifts of f but at least contains good approximations. Therefore we define sampling points $0 \leq \mathbf{t}_1 < \dots < \mathbf{t}_M < \infty$ and choose the dictionary $\mathcal{D} = \{f(\cdot - \mathbf{t}_k)\}_{k=1}^M$. If the sampling distance $\mathbf{t}_{k+1} - \mathbf{t}_k$ is small enough, then we can assume that there exists \mathbf{t}_{j_k} such that $|\mathbf{t}_{j_k} - \tau_k|$ is small and

$$\mathbf{u} = \sum_{k=1}^L x_k f(\cdot - \tau_k) \approx \sum_{k=1}^L \tilde{x}_k f(\cdot - \mathbf{t}_{j_k}).$$

Hence, \mathbf{u} possesses a sparse approximation in \mathcal{D} . We can now discretize equation (6.5) to

$$\begin{pmatrix} \mathbf{u}(t_1) \\ \vdots \\ \mathbf{u}(t_N) \end{pmatrix} = \begin{pmatrix} f(t_1 - \mathbf{t}_1) & \dots & f(t_1 - \mathbf{t}_M) \\ \vdots & \ddots & \vdots \\ f(t_N - \mathbf{t}_1) & \dots & f(t_N - \mathbf{t}_M) \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix}.$$

To avoid notational clutter we will stay with the notation of (6.5) and define

$$\mathbf{u} := \begin{pmatrix} \mathbf{u}(t_1) \\ \vdots \\ \mathbf{u}(t_N) \end{pmatrix}, \quad F := \begin{pmatrix} f(t_1 - \mathbf{t}_1) & \dots & f(t_1 - \mathbf{t}_M) \\ \vdots & \ddots & \vdots \\ f(t_N - \mathbf{t}_1) & \dots & f(t_N - \mathbf{t}_M) \end{pmatrix}, \quad X := \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix},$$

where each column $(f(t_1 - \mathbf{t}_k), \dots, f(t_N - \mathbf{t}_k))^T$ of F is the discretization of an element $f(\cdot - \mathbf{t}_k)$ of the dictionary \mathcal{D} . Note that for $N = M$ and $t_k = \mathbf{t}_k$ for all $k = 1, \dots, N$ the matrix F becomes an ordinary discrete convolution matrix. To perform one MP iteration on the data \mathbf{u} with dictionary \mathcal{D} we need to calculate

$$\begin{aligned} \arg \max_{j=1, \dots, M} |\langle \mathbf{u}, f(\cdot - \mathbf{t}_j) \rangle| &= \arg \max_{j=1, \dots, M} \left| \sum_{k=1}^N \mathbf{u}(t_k) f(t_k - \mathbf{t}_j) \right| \\ &= \arg \max |F^T \mathbf{u}|. \end{aligned}$$

Thus the iteration step of MP and OMP can be performed by evaluating a matrix vector multiplication. However, remember that we have designed our ultrasound impulse function in such a way that they have a finite support and hence the matrix F will be sparse containing only $S \cdot M$ non-zero elements, where S is the size of the support of the discretized impulse function. We note that the sampling resolution of $\{\mathbf{t}_k\}_{k=1}^M$ should be at least as high as the resolution of $\{t_k\}_{k=1}^N$ to avoid a loss of information. However, choosing $M \gg N$ will not improve the method as the correlation between $f(\cdot - \mathbf{t}_k)$ and $f(\cdot - \mathbf{t}_{k+1})$ becomes too large and hamper a correct determination of the right dictionary element even with very few noise on the data. Hence we can assume that M is of size $O(N)$ and thus, the matrix vector multiplication has a complexity of $O(N)$. Since all other computations for one iteration step are also bounded by $O(N)$, the complete MP algorithm can be performed in $O(LN)$ computations and is therefore suitable for real time computation.

As a remark, we want to mention that the general discretization shown above allows a sampling grid $\{\mathbf{t}_k\}_{k=1}^M$ for the solution X that is independent of the sampling grid $\{t_k\}_{k=1}^N$ of the given data. Nevertheless, a typical choice would be $\mathbf{t}_k = t_k$ for $k = 1, \dots, N - M_0$ where we use the same grid on both sides and skip the last M_0 elements as the main part of the support of $f(\cdot - \mathbf{t}_k)$ will be outside the sampling interval $[t_1, t_N]$ and thus does not allow a good correlation analysis, see Figure 6.1. If not mentioned otherwise, this will be the grid used in all numerical examples in this work.

Until now, we have assumed that the number of non-zero elements L of X is known. In practice, the more common case will be that only an upper bound L can be determined but the exact number is unknown. Consider for example a ToFD-measurement. Here, each column of the data might contain a lateral wave, a back wall reflection and not more than 1 – 3 defect signals. Thus $L = 5$ might be a suitable upper bound. But as most columns will only contain the lateral wave and the back wall echo, the MP algorithm will reconstruct at least 3 signals that were caused by noise on the data. This demonstrates the need for a stopping criterion that involves information of the data. We want to suggest two heuristic ideas that have provided good results in numerical experiments.

- The first stopping criterion is based on the norm of the residual r_k . If the norm $\|r_k\|_p$ is smaller than a chosen threshold, then the information contained in the residual is low and it might already only contain noise. As the ultrasound impulse

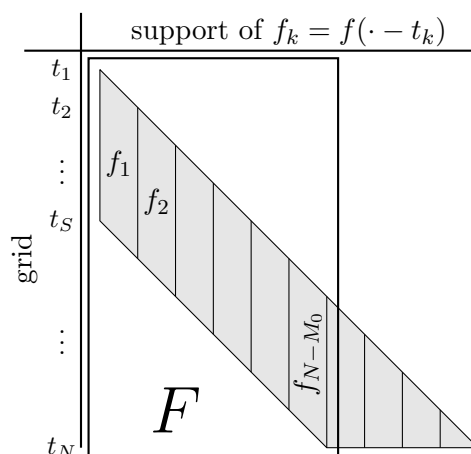


Fig. 6.1: Illustration of F for an equidistant discretization.

has a sparse support and mainly consists of a few high values before it decreases, one may choose $p \gg 1$ up to $p = \infty$ to prefer high values in r_k . Furthermore, because the input data differs strongly depending on the material and the probe set-up, we chose the threshold relative to the norm of \mathbf{u} such that the stopping criterion automatically adapts to the input data. The algorithm will stop if

$$\|r_k\|_p < \varepsilon \|\mathbf{u}\|_2$$

holds for a chosen ε . Note that we use the 2-norm of \mathbf{u} because the original signal may contain multiple reflections and diffractions that are less emphasised in p -norms with large p .

- As a second stopping criterion we consider the value x_k itself. Remember that $|x_k| = \max |F^T \mathbf{u}|$ is the maximum correlation between the residual and our dictionary elements. If this correlation becomes small, the residual may only contain signals that do not fit well into our dictionary, as e.g. noise. As mentioned above, we can choose a threshold relative to the input data and let the MP stop if

$$|x_k| < \varepsilon \|\mathbf{u}\|_2.$$

One advantage of this stopping criterion is that the value x_k is already given by the algorithm itself and hence it does not require the calculation of a p -norm as the criterion shown above does.

We suggest a further adaption of MP and OMP to ultrasonic non-destructive testing data that becomes clear if we remember the model designed in Chapter 4. A basic property of ultrasound we discussed was that its amplitude decreases in time by effects like absorption. It follows that diffractions caused by a defect will have a lower amplitude the longer the distance between probes and defect is. However, although the amplitude is

lower for a far distance, the information carried about the defect boundary is the same. Thus we consider not to choose the dictionary element with the greatest correlation $\arg \max |F^T \mathbf{u}|$ as this will prefer signals that come early in time, but to apply a weighted greatest correlation $\arg \max |WF^T \mathbf{u}|$. Here $W = \text{diag}(w) \in \mathbb{R}^{M \times M}$ is a weight-matrix with the entries $w \in \mathbb{R}^M$ on the diagonal. To compensate the decrease of the amplitude in time we choose w with components of increasing magnitude. Furthermore, one may even choose particular weights very low to e.g. suppress unwanted signals corresponding to the back wall echo or the lateral wave.

All considerations above apply for MP as well as for OMP. To decide which algorithm is the best for our sparse deconvolution problem, we have to bring the differences between MP and OMP back in mind. As we have seen, both algorithms only differ in the calculation of the coefficients x_k . Here OMP solves a minimization problem to deduct the correlation between the active dictionary elements. Although this makes the method more accurate, solving a minimization problem in each iteration is a time consuming process. In order to use the accuracy of OMP without a drastic increase of computation time we propose a MP-OMP-hybrid method, where the coefficients are determined by the OMP minimization problem only if the correlation of the newly added dictionary element and the already active elements is too strong. Otherwise the coefficient is directly determined via the MP iteration.

Combining all considerations above we obtain the following hybrid orthogonal matching pursuit (HOMP) algorithm adapted to ultrasonic non-destructive testing.

Algorithm 6.1.1 (HOMP for NDT data):

Let the input data \mathbf{u} and the function f be given and let F_k denote the matrix F restricted to the columns j_1, \dots, j_k , where $F_0 := 0$. We perform the following HOMP algorithm to solve the sparse deconvolution problem:

1. Choose a sampling grid $\{\mathbf{t}_k\}_{k=1}^M$ and calculate $F = (f(t_j - \mathbf{t}_k))_{j,k=1}^{N,M}$. Set $W = \text{diag}(w)$, $L := 5$, $k := 1$, $p \gg 1$ and $\varepsilon_0, \varepsilon_1, \varepsilon_2 > 0$. Initialize $X := 0 \in \mathbb{R}^M$ and $r_1 := \mathbf{u}$.
2. Determine $j_k := \arg \max |WF^T r_k|$.
3. Calculate $E := \|F_{k-1} \cdot (f(t_i - \mathbf{t}_{j_k}))_{i=1}^N\|_2$.
4. If $E < \varepsilon_0$ set $x_{j_k} := (F^T r_k)_{j_k}$ else set $(x_{j_i})_{i=1}^k := \arg \min_{y \in \mathbb{R}^k} \|\mathbf{u} - F_k y\|_2$.
5. Update $r_{k+1} := \mathbf{u} - FX$ and $k := k + 1$.
6. Repeat 2 – 4 until $k > L$, $\|r_k\|_p < \varepsilon_1 \|\mathbf{u}\|_2$ or $|x_{j_k}| < \varepsilon_2 \|\mathbf{u}\|_2$.
7. Return X .

Here E calculates the correlation between the newly chosen dictionary element and those which are already active. So far we did not analyse the approximation results and the stability of MP based algorithms. There are already many results known as e.g. shown in [77]. Unfortunately, most of these results require a dictionary where the correlation between two elements is bounded away from 1. Thus, these results do not apply in our case as the dictionary elements are strongly correlated especially for neighbouring shifts.

However, as we will see in the numerical examples in Chapter 9, the method provides good results even for data with strong noise. Nonetheless, in the following subsection we briefly want to discuss some basic examples for MP that will give an idea of the problems that may arise in real data computation.

6.1.3 Some Comments on Approximation Properties of MP

Although MP-like algorithms are in many cases successfully applied to reconstruct a sparse solution, the corresponding theory does not cover all applications. Especially for dictionaries with highly correlating elements the behaviour of MP is somehow unclear. Hence we will consider a few examples in this subsection to get an idea of the answer to the following questions: How does noise influence MP? How does MP act on signals containing highly correlated dictionary elements (two close by non-zero elements in X)? Does a false reconstruction at least approximate the right solution?

First, we want to analyse the behaviour of MP when noise is added to the data. Therefore we consider the functions $f_k = f(\cdot - k\delta_t)$ with $k \in \mathbb{Z}$, $\delta_t \in \mathbb{R}_+$, where $\|f\|_{\mathcal{H}} = 1$ and f has finite support. For simplicity we use the infinite dictionary $\mathcal{D} = \{f_k \mid k \in \mathbb{Z}\}$ to avoid boundary effects. We want to reconstruct the 1-term approximation of $\mathbf{u} = af_{k'}$ with $k' \in \mathbb{Z}$. If there is no noise on the data \mathbf{u} , then the MP-algorithm will surely find the exact solution as

$$k' = \arg \max_{k \in \mathbb{Z}} |\langle \mathbf{u}, f_k \rangle|$$

and

$$a = \langle \mathbf{u}, f_{k'} \rangle.$$

However, let us now consider noisy data

$$\mathbf{u} = af_{k'} + \sum_{k=-\infty}^{\infty} \varepsilon_k f_k$$

with Gaussian random variables ε_k with variance σ^2 . First, we note that

$$\langle \mathbf{u}, f_{k'} \rangle = a + \sum_{k=-\infty}^{\infty} \varepsilon_k \langle f_k, f_{k'} \rangle$$

will most unlikely be equal to a and thus MP cannot recover the exact amplitude. But furthermore because of

$$\langle \mathbf{u}, f_l \rangle = a \langle f_{k'}, f_l \rangle + \sum_{k=-\infty}^{\infty} \varepsilon_k \langle f_k, f_l \rangle = a \langle f_{k'}, f_l \rangle + \varepsilon$$

with a Gaussian random variable ε with variance $\sum_{k=-\infty}^{\infty} \langle f_k, f_l \rangle^2 \sigma^2 := C(\delta_t) \sigma^2$, we obtain

$$\langle \mathbf{u}, f_{k'} \rangle - \langle \mathbf{u}, f_l \rangle = a(1 - \langle f_{k'}, f_l \rangle) + \varepsilon' \quad (6.6)$$

with a Gaussian random variable ε' with variance $2C(\delta_t) \sigma^2$. If we now consider the case where l is close to k' , then the correlation $\langle f_{k'}, f_l \rangle$ may be high and thus for σ large enough (or δ_t small enough) it is unclear whether $\langle \mathbf{u}, f_{k'} \rangle > \langle \mathbf{u}, f_l \rangle$ or not. Hence the MP-algorithm may choose the wrong dictionary element due to the noise. From this observation, it follows that the MP-algorithm has a restricted resolution depending on the noise level. Note that $C(\delta_t)$ decreases for increasing step size δ_t and as a consequence the method seems to be more stable for large step sizes. However, this does only hold if \mathbf{u} is the combination of dictionary elements. In the general case where \mathbf{u} is unknown, one has to use a suitable small step size to ensure that \mathbf{u} can be approximated by a combination of the dictionary elements.

For the next example we consider the dictionary $\mathcal{D} = \{f_1, f_2, f_3\}$ with

$$\begin{aligned} f_1(x) &= \begin{cases} \sin(x) & 0 \leq x \leq 2\pi \\ 0 & \text{else} \end{cases}, \\ f_2(x) &= \begin{cases} \sin(x - \alpha) & \alpha \leq x \leq 2\pi + \alpha \\ 0 & \text{else} \end{cases}, \\ f_3(x) &= \begin{cases} \sin(x - 2\alpha) & 2\alpha \leq x \leq 2\pi + 2\alpha \\ 0 & \text{else} \end{cases} \end{aligned}$$

with $\alpha < \frac{\pi}{4}$. This is a very simple dictionary of shifted and truncated sine functions as ultrasound impulses. Note that the functions are not normalized but this can be neglected for the following considerations as all three functions have the same norm. We obtain

$$\begin{aligned} \langle f_2, f_3 \rangle = \langle f_1, f_2 \rangle &= \int_{\alpha}^{2\pi} \sin(x) \sin(x - \alpha) dx = \frac{2\pi - \alpha}{2} \cos(\alpha) + \frac{\sin \alpha}{2} > 0 \\ \langle f_1, f_3 \rangle &= (\pi - \alpha) \cos(2\alpha) + \frac{\sin(2\alpha)}{2} > 0. \end{aligned}$$

Consider the signal $\mathbf{u} = f_1 + f_3$. It follows

$$\begin{aligned} \langle \mathbf{u}, f_1 \rangle &= \pi + \langle f_1, f_3 \rangle \\ \langle \mathbf{u}, f_2 \rangle &= \langle f_1, f_2 \rangle + \langle f_3, f_2 \rangle \\ \langle \mathbf{u}, f_3 \rangle &= \langle f_1, f_3 \rangle + \pi. \end{aligned}$$

To ensure that the MP-algorithm chooses the right dictionary element, namely f_1 or f_3 , $\langle \mathbf{u}, f_1 \rangle > \langle \mathbf{u}, f_2 \rangle$ has to hold. But for e.g. $\alpha = \frac{\pi}{6}$, we obtain

$$\begin{aligned} \langle \mathbf{u}, f_1 \rangle - \langle \mathbf{u}, f_2 \rangle &= \pi + (\pi - \alpha) \cos(2\alpha) + \frac{\sin(2\alpha)}{2} - (2\pi - \alpha) \cos(\alpha) - \sin(\alpha) \\ &\approx -0.6043 \\ &< 0 \end{aligned}$$

and thus $\langle \mathbf{u}, f_1 \rangle < \langle \mathbf{u}, f_2 \rangle$, i.e. the MP-algorithm will choose the wrong dictionary element for small α . We note that the method tries to approximate the signal \mathbf{u} with the element that fits best and, caused by the strong correlation of f_2 with f_1 and f_3 , it does not detect two active dictionary functions but approximates them with a function that is right in the middle.

We have now seen some limitations of MP-like algorithms. While one can theoretically choose a huge dictionary \mathcal{D} to provide a large number of elements MP can choose, the resolution will not benefit as it is bounded by the noise and the correlation of the elements. Especially if the signal \mathbf{u} consists of several strongly correlated elements the algorithm will most likely not find the right elements. Note that the above examples do not only hold for MP itself but for OMP and other modifications as well, as they all choose the next active element in the same way due to the highest correlation. However, OMP, the here presented HOMP for NDT data, and also other techniques may prove to deliver better results for a high number of iterations as they also try to balance the error from previous iterations, while MP does not. But although MP-like algorithms might not always find the best sparse solution of the data \mathbf{u} , considering (6.3) it is likely that they will find a suitable approximation.

6.2 Blind Deconvolution

In the last section we have introduced a method to solve the sparse deconvolution problem $\mathbf{u} = f * X$ if f is known. Yet, the more common case in practise is that the wave impulse f is not known. This means, we have to estimate both, the sparse solution X and the pulse function f from the given data $\mathbf{u} = f * X$. These so-called blind deconvolution problems are of special interest and many approaches to solve it have been developed, see e.g. [74, 91], the adaptive approaches in [92, 93] based on minimum entropy evaluation, [94, 95] considering order statistics or [96, 97] using wavelet regularisation. We also refer to similar methods in [98, 99] that are directly applied to b-scans of ultrasonic data.

However, to receive fast and stable algorithms for blind deconvolution of ultrasonic data one really needs to exploit the sparsity of X and the special structure of the wave impulse f as these are powerful constraints. Therefore, we consider the data $\mathbf{u} \in \mathbb{R}^N$ sampled at $0 \leq t_1 < t_2 < \dots < t_N$. As f is an ultrasound impulse, it can be modelled as a function $f = f_{\mathbf{p}}$ depending on parameters \mathbf{p} , as e.g. the Gabor function (4.13) with $\mathbf{p} = (\rho, \psi, \phi)$

or the raised cosine (4.14) with $\mathbf{p} = (N, \psi)$. For a sampling $0 \leq \tau_1 < \tau_2 < \dots < \tau_M$, define $f_k = (f_{\mathbf{p}}(t_1 - \tau_k), \dots, f_{\mathbf{p}}(t_N - \tau_k))$. Using the sparsity of X we can now rewrite $\mathbf{u} = f * X$ as

$$\mathbf{u} = \sum_{j=1}^S x_{v_j} f_{v_j} \quad (6.7)$$

where v_j is the index of the j -th non-zero element of X and S is the number of non-zero elements. Besides S , v_j and x_{v_j} , $j = 1, \dots, S$, we now also have to estimate the parameter set \mathbf{p} defining the function $f_{\mathbf{p}}$ that generates the dictionary. In this work we propose the following approach:

1. Calculate a starting guess of $\tilde{\mathbf{p}} \approx \mathbf{p}$.
2. Use the HOMP algorithm (6.1.1) to derive S , $\{v_j\}_{j=1}^S$ and $\{x_{v_j}\}_{j=1}^S$.
3. Set $\tilde{\mathbf{p}} := \arg \min_{\mathbf{p}} \|\mathbf{u} - \sum_{j=1}^S x_{v_j} f_{v_j}\|_2$.
4. Iterate 2-3 until a stopping criterion holds.

We will present an easy way to calculate a very good starting guess $\tilde{\mathbf{p}}$ from the given data in the following subsection. Although there is no guaranty for global convergence of step 2 and 3, this choice of $\tilde{\mathbf{p}}$ will enable reasonable results even in the very first iterations. Note that step 3 is a minimization problem only in the parameter vector \mathbf{p} and thus of low dimension. Its solution can be approximated with fast converging algorithms as e.g. Newton iteration. We will discuss this approach later.

Furthermore, note that our model is a simplified version of the model presented in [74, 100]. There one uses the functions $f_k = (f_{\mathbf{p}_1}(t_1 - \tau_k), \dots, f_{\mathbf{p}_N}(t_N - \tau_k))$, meaning that each shift function has its own parameter vector. This general approach might be useful if one considers inhomogeneous material where the structure of the incoming wave impulses are strongly influenced by the material. The additional degree of freedom is traded against a higher complexity of the problem as one now also has to estimate $S \# \mathbf{p}$ parameters. As a drawback, the number of non-zero elements S has to be known in the general model, see [74]. Furthermore, the used algorithms, namely expectation maximization (EM) or space alternating generalized EM (SAGE) [101], might either be slow [102] or unstable [103]. However, we assume homogeneous material and thus it is a suitable assumption that the measured wave impulses have (approximatively) the same structure.

6.2.1 Calculation of the Starting Parameters

First, let us consider how one obtains a suitable starting guess $\tilde{\mathbf{p}}$. The easiest way to do so, is to use information given by the testing set-up as e.g. the a-priori known frequency

of the emitter pulse. But although the sent out ultrasound impulse is given by the emitters properties, the parameters might change slightly due to the attenuation in the material. Thus it is advisable to also use the information given by the data \mathbf{u} itself as it already includes this change of the pulse shape. To obtain a reliable approximation we consider several data vectors $\mathbf{u}_1, \dots, \mathbf{u}_Z$, $Z \in \mathbb{N}$ produced with the same set-up. This procedure does not cause extra costs in most cases since the NDT routine will perform multiple measurements anyway. Our method to derive a suitable $\tilde{\mathbf{p}}$ will be more accurate the more measurements are taken, i.e. the larger Z is chosen. However, we only want to calculate a starting guess, and the parameter vector will be updated in each iteration. Thus it is not necessary to choose Z very large.

Given the data vectors $\mathbf{u}_1, \dots, \mathbf{u}_Z \in \mathbb{R}^N$, $Z \in \mathbb{N}$ that have been sampled with the same set-up at time $0 \leq t_1 < \dots < t_N$, we concentrate on a signal that is contained in every measurement \mathbf{u}_k , $k = 1, \dots, Z$, namely the back wall echo. Here we do not consider surface reflections or the lateral wave as these signals are less reliable and can be removed by cutting of the first elements of each vector. It is most likely that in each vector \mathbf{u}_k the reflection caused by the back wall is represented by the same shift of the function $f_{\mathbf{p}}$ and has the same amplitude. Thus, if we sum up all data vectors, the back wall echo will remain due to constructive interference while all other signals will cause negligibly small amplitudes. We obtain

$$\sum_{k=1}^Z \mathbf{u}_k \approx x (f_{\mathbf{p}}(t_j - \tilde{t}))_{j=1}^N \quad (6.8)$$

with an amplitude $x \in \mathbb{R}$. The unknown shift \tilde{t} can be calculated easily as we know that $f_{\mathbf{p}}(\cdot - \tilde{t})$ has its essential support on an interval that starts at time \tilde{t} . Hence for a chosen $\varepsilon > 0$ we set

$$\tilde{t} := \min \left\{ t_j \mid \sum_{k=1}^Z \mathbf{u}_k(t_j) > \varepsilon \right\} \text{ and } \tilde{t}_0 := \max \left\{ t_j \mid \sum_{k=1}^Z \mathbf{u}_k(t_j) > \varepsilon \right\},$$

Furthermore we divide (6.8) by its maximum absolute value to normalize it. It follows that

$$\tilde{\mathbf{f}} := \left(\frac{\sum_{k=1}^Z \mathbf{u}_k(t_l)}{\max_{j=1, \dots, N} \left| \sum_{k=1}^Z \mathbf{u}_k(t_j) \right|} \right)_{\tilde{t} \leq t_l \leq \tilde{t}_0} \approx \frac{(f_{\mathbf{p}}(t_l - \tilde{t}))_{\tilde{t} \leq t_l \leq \tilde{t}_0}}{\|f_{\mathbf{p}}\|_{\infty}}. \quad (6.9)$$

A very good starting guess $\tilde{\mathbf{p}}$ can now be found by solving the problem

$$\tilde{\mathbf{p}} = \arg \min_{\mathbf{p}} \left\| \tilde{\mathbf{f}} - \frac{(f_{\mathbf{p}}(t_l - \tilde{t}))_{\tilde{t} \leq t_l \leq \tilde{t}_0}}{\|f_{\mathbf{p}}\|_{\infty}} \right\|_2. \quad (6.10)$$

Figure 6.2 illustrates this technique of deriving a starting guess for the case that f is a Gabor function as given in (4.13). The left column shows the summation of several a-scans (6.8), where the first two rows were obtained from a ToFD measurement while the last row corresponds to a wall thickness measurement. The second column illustrates the normalization and support restriction (6.9), and in the last column one can see the solution of the minimization problem (6.10). The method reconstructs the starting guess

$$\begin{aligned} \tilde{\mathbf{p}} &= (6.8486, 14.685, -2.0836) && \text{(top)} \\ \tilde{\mathbf{p}} &= (30.0, 28.039, 3.0867) && \text{(middle)} \\ \tilde{\mathbf{p}} &= (45.0, 35.448, 1.5708) && \text{(bottom)} \end{aligned}$$

given in (MHz², Mhz, ·).

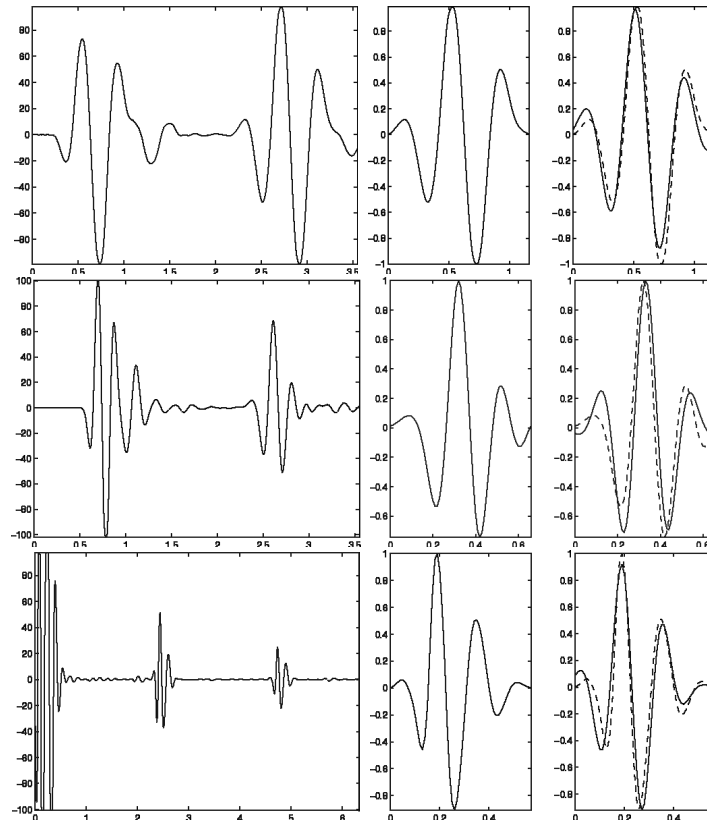


Fig. 6.2: Calculation of the starting guess $\tilde{\mathbf{p}}$, left: the summation of several a-scans, middle: normalization and restriction to the essential support of the back wall echo, right: approximation by a Gabor impulse; time in microseconds.

6.2.2 Parameter Update by Newton-Iteration

Now it is clear how we can find suitable starting parameters $\tilde{\mathbf{p}}$, and hence a first approximation of X can be found with the HOMP algorithm (6.1.1). To further improve the solution, we want to apply alternately HOMP for sparse deconvolution and a Newton algorithm for parameter updating. The HOMP algorithm has already been discussed in the first section of this Chapter. Thus we will now consider the minimization problem

$$\tilde{\mathbf{p}} = \arg \min_{\mathbf{p}} \left\| \mathbf{u} - \sum_{j=1}^S x_{v_j} f_{v_j} \right\|_2 \quad (6.11)$$

and how it can be solved with a Newton-Iteration approach. Therefore we define the function

$$\mathfrak{F}(\mathbf{p}) = \sum_{j=1}^S x_{v_j} f_{v_j}(\mathbf{p}),$$

where we explicitly denote the dependence of $f_{v_j}(\mathbf{p})$ on \mathbf{p} . The minimization (6.11) can now be written as $\tilde{\mathbf{p}} = \arg \min \|\mathbf{u} - \mathfrak{F}(\mathbf{p})\|_2$. We will solve this problem by linearising the operator \mathfrak{F} and performing several Newton-iterations to calculate the minimum of $\|\mathbf{u} - \mathfrak{F}(\mathbf{p})\|_2$ with respect to \mathbf{p} . We obtain the following algorithm:

1. For a given starting parameter $\tilde{\mathbf{p}}$ set $\mathbf{p}_1 := \tilde{\mathbf{p}}$, $k := 1$ and choose $\varepsilon > 0$.
2. Solve $\mathfrak{F}' \Delta_{\mathbf{p}} = \mathbf{u} - \mathfrak{F}(\mathbf{p}_k)$ in a least square sense e.g. with the Pseudo inverse. Here $\mathfrak{F}' \in \mathbb{R}^{N \times \#\mathbf{p}}$ denotes the Jacobian matrix of \mathfrak{F} .
3. Update $\mathbf{p}_{k+1} := \mathbf{p}_k + \Delta_{\mathbf{p}}$ and $k := k + 1$.
4. Iterate 2-3 until the step size $\|\Delta_{\mathbf{p}}\|_2$ satisfies $\|\Delta_{\mathbf{p}}\|_2 < \varepsilon$.

To perform this algorithm we need to calculate the Jacobian matrix \mathfrak{F}' . We will do this analytically for the case of Gabor impulses with the parameters $\mathbf{p} = (\rho, \psi, \phi)$, i.e. for

$$f_{\mathbf{p}}(t) = e^{-\rho t^2} \cos(\psi t + \phi),$$

where we explicitly do not normalize the function f as the normalization factor will depend on the parameters \mathbf{p} and thus an analytical differentiation will get complicated. Because of

$$\begin{aligned} \frac{d}{d\rho} f &= -t^2 e^{-\rho t^2} \cos(\psi t + \phi) \\ \frac{d}{d\psi} f &= -t e^{-\rho t^2} \sin(\psi t + \phi) \\ \frac{d}{d\phi} f &= -e^{-\rho t^2} \sin(\psi t + \phi) \end{aligned}$$

we obtain $\mathfrak{F}' \in \mathbb{R}^{N \times 3}$ with

$$\mathfrak{F}' = \left(\frac{d}{d\rho} \mathfrak{F}, \frac{d}{d\psi} \mathfrak{F}, \frac{d}{d\phi} \mathfrak{F} \right) = \begin{pmatrix} -\sum_{j=1}^S x_{v_j} (t_k - \mathbf{t}_{v_j})^2 e^{-\rho(t_k - \mathbf{t}_{v_j})^2} \cos(\psi(t_k - \mathbf{t}_{v_j}) + \phi) \\ -\sum_{j=1}^S x_{v_j} (t_k - \mathbf{t}_{v_j}) e^{-\rho(t_k - \mathbf{t}_{v_j})^2} \sin(\psi(t_k - \mathbf{t}_{v_j}) + \phi) \\ -\sum_{j=1}^S x_{v_j} e^{-\rho(t_k - \mathbf{t}_{v_j})^2} \sin(\psi(t_k - \mathbf{t}_{v_j}) + \phi) \end{pmatrix}_{k \leq N}^T.$$

Unfortunately, as we consider a function f that is not normalized, the update of the parameters \mathbf{p} implies a possibly considerable change of the norm of f . This leads to a highly unstable method since the amplitudes $\{x_{v_j}\}_{j=1}^S$ are optimized with respect to the norm of $f_{\tilde{\mathbf{p}}}$ with the parameters $\tilde{\mathbf{p}}$ obtained beforehand. In order to counter this problem we will update not only the parameters in each Newton-Iteration but also the amplitudes $\{x_{v_j}\}_{j=1}^S$ to allow an adjustment of the amplitudes to the changing norm of the wave function. Note that we do not change the positions $\{v_j\}_{j=1}^S$ of the non-zero elements in X . This modifies the Jacobian matrix \mathfrak{F}' as it is also calculated with respect to the amplitudes and thus $\mathfrak{F}' \in \mathbb{R}^{N \times S \# \mathbf{p}}$ is given by

$$\mathfrak{F}' = \begin{pmatrix} -\sum_{j=1}^S x_{v_j} (t_k - \mathbf{t}_{v_j})^2 e^{-\rho(t_k - \mathbf{t}_{v_j})^2} \cos(\psi(t_k - \mathbf{t}_{v_j}) + \phi) \\ -\sum_{j=1}^S x_{v_j} (t_k - \mathbf{t}_{v_j}) e^{-\rho(t_k - \mathbf{t}_{v_j})^2} \sin(\psi(t_k - \mathbf{t}_{v_j}) + \phi) \\ -\sum_{j=1}^S x_{v_j} e^{-\rho(t_k - \mathbf{t}_{v_j})^2} \sin(\psi(t_k - \mathbf{t}_{v_j}) + \phi) \\ e^{-\rho(t_k - \mathbf{t}_{v_1})^2} \cos(\psi(t_k - \mathbf{t}_{v_1}) + \phi) \\ \vdots \\ e^{-\rho(t_k - \mathbf{t}_{v_S})^2} \cos(\psi(t_k - \mathbf{t}_{v_S}) + \phi) \end{pmatrix}_{k=1, \dots, N}^T.$$

The numerical results presented in Chapter 9 show that the Newton-iteration performed with this Jacobian matrix is stable and converges fast even for significant changes in the norm of $f_{\mathbf{p}}$. It emerges to be sufficient to use only very few iteration steps, namely no more than about 10, to obtain reasonable results.

6.3 Summary

Let us summarize the results achieved in this chapter. In the first step of our reconstruction method we had to solve the sparse blind deconvolution problem $\mathbf{u} = f * X$. In Section 6.1 we discussed the MP-based method HOMP (6.1.1) to reconstruct the sparse solution X if f is known. The great advantage of this algorithm is its fast and stable implementation. However, the resolution of the recovered solution is limited by

the noise level and HOMP may fail to reconstruct all non-zero elements if some of them are strongly correlated. Moreover, in practical applications the function f is often unknown. Therefore we introduced an approach to solve the blind deconvolution problem by alternately using HOMP for sparse deconvolution and Newton-iterations to update the function f_p in Section 6.2. We used the Gabor impulse as a suitable model for f_p with three parameters. We also provided a technique to derive a suitable starting guess as the Newton method only converges locally. Altogether the following algorithm has been developed:

Algorithm 6.3.1 (Newton-HOMP for sparse blind deconvolution of NDT data):

Given the input data u perform the following operations:

1. Calculate the starting parameters \tilde{p} as described in Subsection 6.2.1.
2. Reconstruct X using the HOMP algorithm (6.1.1).
3.
 - a) Solve $\tilde{\mathfrak{F}}' \Delta_p = u - \tilde{\mathfrak{F}}(\tilde{p})$ in a least square sense.
 - b) Update $\tilde{p} := \tilde{p} + \Delta_p$.
 - c) Iterate a) and b) for $\mathfrak{N} \leq 10$ times.
4. Iterate step 2 and 3 until the non-zero elements of X do not change any more.
5. Return X .

Figure 6.3 illustrates the numerical results of this algorithm for real NDT data. For a detailed numerical discussion we refer to Chapter 9 of this work. Now we will move on to the next chapter where we analyse the second step of our inversion method.

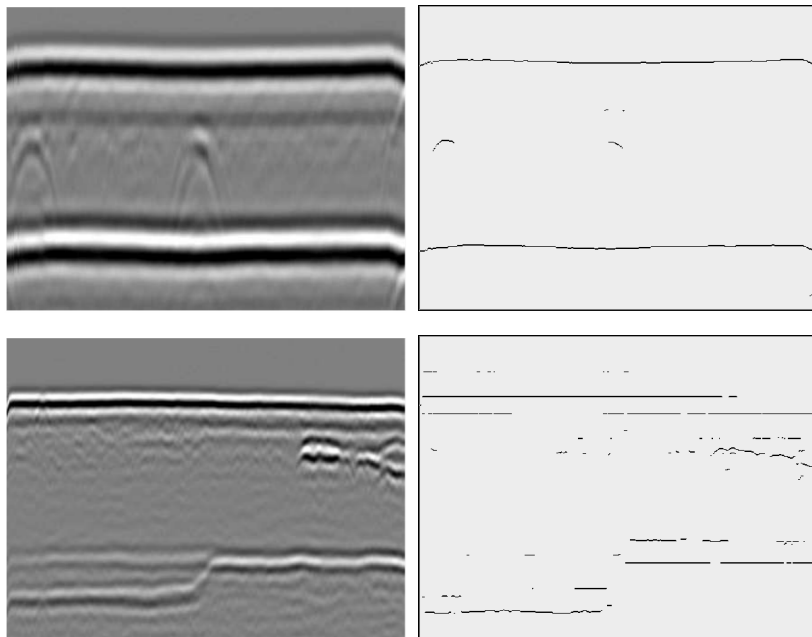


Fig. 6.3: Sparse blind deconvolution with Algorithm 6.3.1 for real NDT data, left: original data, right: non-zero elements (black) after deconvolution.

7 Inversion 2 - Clustering

In the last Chapter we discussed a method to deconvolve and denoise the given data \mathbf{u} . Therefore a sparse blind deconvolution problem $\mathbf{u} = f * X$ had to be solved to reconstruct the large amplitudes of X and to suppress the small noise. We will now use the obtained information to group the non-zero elements of X , where all elements in one group will most likely be caused by the same defect in the material. Note that for simplicity and motivated by Figure 6.3 we will from now on refer to the non-zero elements of X as *peaks*.

While the deconvolution can be performed independently on each a-scan, i.e. on each column of the data matrix \mathbf{u} , here it is essential to not only consider the non-zero elements of one a-scan but also use the information given by neighbouring a-scans. A defect that causes a peak in one of the a-scans will probably cause more peaks for nearby probe positions. Thus, for each a-scan at its own the peak does not provide suitable information but considering several a-scans the peaks will accumulate in a region and form a cluster. We will use a clustering algorithm to extract this information and reconstruct the number of defects as well as the peaks caused by each of it. Therefore, we will only use the positions of the in step 1 reconstructed non-zero elements since the amplitudes are difficult to interpret even without noise because they are highly dependent on many material and defect properties. The problem can be stated as follows:

Given the pairs $\{(t_j, h_j)\}_{j=1}^{S_N}$, $S_N = O(NS)$ extracted out of N a-scans for which

$$\sum_{k=1}^K \int_{\partial D_k \cap E_{h_j, t_j}} a(t_j, y, h_j) dy \neq 0$$

holds, we want to find the number of defects $K \in \mathbb{N}$ and $\{(t_{j,k}, h_{j,k})\}_{j,k=1}^{J_k, K}$ such that

$$\int_{\partial D_k \cap E_{h_{j,k}, t_{j,k}}} a(t_{j,k}, y, h_{j,k}) dy \neq 0$$

holds for all $k = 1, \dots, K$ and $j = 1, \dots, J_k$. Note that as a result of the sparse deconvolution performed in step 1, the input data $\{(t_j, h_j)\}_{j=1}^{S_N}$ given for step 2 only consists of $S_N = O(SN) = O(N)$ elements and thus is independent of the length of the columns of \mathbf{u} . In this way, the sparse deconvolution does not only extract important

information out of the measurement data but does also decrease the numerical effort for the following steps drastically.

The above formulated problem is a typical application of cluster analysis. A given number of data points, here positions (t_j, h_j) , has to be sorted into different groups, where the number of groups is also unknown. Thus, we will shortly introduce some basic ideas of cluster analysis in the next subsection and discuss our application afterwards, where we will first consider the simple case of defects whose caused peaks are separated in space and later the case where the clusters may overlap. For that matter it will prove to be useful to use information about the data provided e.g. by our model. This information will be crucial for a good and stable clustering algorithm.

7.1 Cluster Analysis

In this section, we want to give a very brief introduction to cluster analysis. Note that only the main ideas will be discussed and we will not introduce any method in detail. This will be done in the following sections when we apply the algorithms to our problem. Also the theoretical considerations are kept short as they are often depending on the methods itself. For a more detailed overview of cluster analysis, we refer to [104–106].

Clustering algorithms aim to classify a given amount of data into several clusters of *similar* objects. Thereby a lot of different clustering techniques have been developed. We want to shortly summarize the most important differences that have to be considered before choosing a clustering algorithm for a special application.

At first, we need to clarify the meaning of *cluster* to understand the principles of clustering algorithms. Besides several requirements that depend on the application, the clusters should have the following two properties [104]:

- Elements in the same cluster should be related to each other / to the cluster.
- Elements of two different clusters should be unrelated to each other / to the cluster.

Therefore, a relation measure has to be defined depending on the application. Consider e.g. a number of points $\{x_k\}_{k=1}^K \in \mathbb{R}^n$. We may say that two elements x_i, x_j are related, if $\|x_i - x_j\|_2$ is small (e.g. below a certain threshold) and are unrelated if $\|x_i - x_j\|_2$ is large. A clustering algorithm then tries to find clusters of points that are near to each other while the distance between two clusters is large. On this way a number of clusters is created that contain related objects with respect to the relation measure. The intensity of relation within one cluster can be chosen arbitrarily. As an example, the Complete-Linkage algorithm [104] creates clusters in which each element is related to each other element of the cluster while the Single-Linkage algorithm [104] only requires each element to be related to at least one other element. Furthermore, some algorithms consider a representative element for each cluster and measure the relation between an element of the data and the representative [107].

We can now give a general description that applies to most clustering algorithms. Given a data set, perform the following procedure:

1. Choose a starting set of clusters, where e.g. each element is a cluster on its own, all elements are in one cluster or, if the number of clusters is known, all elements are randomly assigned to a cluster.
2. Calculate all values of the relation measure.
3. Update the clusters by e.g. merging clusters, separating clusters or reassigning elements.
4. Iterate 2-3 until the clusters do not change any more or another stopping criterion holds.

Thereby, the detailed procedure of all steps is dependent on the exact methods and will be discussed in the following sections for the algorithms we use in this work. To give an example, we refer to one of the most used clustering algorithms, the k-means [104, 107, 108].

However, the number of algorithms one can apply might be limited by the application. As an example, some clustering algorithms as e.g. the k-means do explicitly require a-priori knowledge about the number of clusters which is not always known. Moreover, some algorithms produce overlapping clusters, meaning that they allow an element to be member of more than one cluster. This may be useful or not depending on the application and interpretation of the clusters. Considering all this, to find a suitable clustering algorithm for a given application, we need to answer the following questions:

- Which function does best measure the relation between our data elements?
- Does it make sense to allow overlapping clusters?
- Do we know the number of clusters?

This will limit the number of algorithms we can use. In order to find the best algorithm we should furthermore take into account the runtime, the relation intensity requested in each cluster and other properties of the different methods. In the following two sections we will consider these points, first for NDT data where the signals of different defects do not intersect and later where they do. For both cases we will present an algorithm that fits the requirements and thus can be taken as the second step of our inversion method.

7.2 Clustering Algorithms for Separated Defect Peaks

Let us analyse how to apply clustering algorithms to reconstruct the number of defects and to classify each peak by its source. Thereby, we first consider the more simple case,

where peaks from different defects are separated. This means, for (\mathbf{t}_j, h_j) , $(\mathbf{t}_{j'}, h_{j'})$ with $j \neq j'$, $k \neq k'$ and

$$\int_{\partial D_k \cap E_{h_j, \mathbf{t}_j}} a(\mathbf{t}_j, y, h_j) dy \neq 0 \qquad \int_{\partial D_{k'} \cap E_{h_{j'}, \mathbf{t}_{j'}}} a(\mathbf{t}_{j'}, y, h_{j'}) dy \neq 0$$

we assume that

$$\sqrt{\gamma(\mathbf{t}_j - \mathbf{t}_{j'})^2 + \frac{1}{2}((x_E(h_j) - x_E(h_{j'}))^2 + (x_R(h_j) - x_R(h_{j'}))^2)} > \varepsilon \quad (7.1)$$

for some fixed $\varepsilon > 0$. Here, $\gamma > 0$ in (7.1) is a weight parameter that is necessary as we are mixing a time and a distance variable. For an equal rate we choose $\gamma = c^2$ with the ultrasound speed c . It follows directly that if (7.1) does not hold, then (\mathbf{t}_j, h_j) and $(\mathbf{t}_{j'}, h_{j'})$ are caused by the same defect, i.e. $k = k'$. Thus the function

$$d_\gamma(\mathbf{t}_j, h_j, \mathbf{t}_{j'}, h_{j'}) = \sqrt{\gamma(\mathbf{t}_j - \mathbf{t}_{j'})^2 + \frac{1}{2}((x_E(h_j) - x_E(h_{j'}))^2 + (x_R(h_j) - x_R(h_{j'}))^2)} \quad (7.2)$$

will be a good choice as a relation measure.

Let us proceed by considering the results of the last section. We note that, due to the separated defects, no peak will be caused by two defects at once and hence we do not want the clusters to overlap. Furthermore, we do not know the number of clusters as it coincides with the also unknown number of defects. Moreover, as there might be a large distance between the first and the last peak caused by a defect, we do not require one element to be related with all other elements of the cluster. An algorithm that satisfies most of these requirements is the DBSCAN [105, 109, 110] (density based spatial clustering for applications with noise). DBSCAN might produce slightly overlapping clusters but as a big advantage, it is also able to handle noisy data, i.e. data points that should not be classified to any cluster. Although most of the noise in our application data will be eliminated in the first inversion step, this cluster algorithm will remove most of the rest. The idea of DBSCAN is, to analyse the local density at each data point. It distinguishes between *core points* with high density, i.e. points whose ε -neighbourhood contains more than P_{\min} points, and points with low density. The parameters $\varepsilon > 0$ and $P_{\min} \in \mathbb{N}$ have to be defined by the user. Starting at a random point DBSCAN runs through the data until a core point is found. The core point and its ε -neighbourhood form the basis of a new cluster. DBSCAN expands the cluster by adding the ε -neighbourhood of all core points in the cluster and iterates this step until there is no new core point added to the cluster. In this way, a cluster consist of so-called *density-reachable* core points and their ε -neighbourhoods. After one cluster is completely expanded, the algorithms starts searching another core point in the remaining set. If in doing so a point with low density is found it will be labelled as noise but might be later reassigned to a cluster if

found in the ε -neighbourhood of a core point. Hence, DBSCAN only needs to consider each data point once to check if it is either a core point and can directly be assigned to a cluster or not.

Let us now adapt DBSCAN to our application. The choice of ε is given by inequality (7.1). Furthermore, we set the minimum number of points in the ε -neighbourhood for core points P_{\min} depending on the application. For e.g. ToFD we set $P_{\min} = 2$ as for each peak their should be at least one peak in the successive and in the previous a-scan and similarly we choose $P_{\min} = 4$ for wall thickness measurements because here one analyses three dimensional data. We obtain:

Algorithm 7.2.1 (DBSCAN for defect-cluster reconstruction):

Given the pairs $\{(t_j, h_j)\}_{j=1}^{S_N}$, let (7.1) hold for $\varepsilon > 0$ for peaks caused by different defects. Choose $\gamma > 0$, P_{\min} arbitrarily and perform the following algorithm:

Initialize the index set $I_{uv} = \{1, \dots, S_N\}$ of unvisited points and the number of clusters $K = 0$.

Search for a core point by

1. Choose $i \in I_{uv}$ and update $I_{uv} = I_{uv} \setminus \{i\}$.
2. Calculate the indices \mathcal{E} of the points in the ε -neighbourhood of (t_i, h_i) w.r.t. the distance (7.2).
3. If \mathcal{E} contains at least P_{\min} indices, start expanding, else continue with step 1.

Expand the cluster by

1. Set $K = K + 1$ and $C_K = \{i\} \cup \mathcal{E}$.
2. For all core points in C_K , add their ε -neighbourhood to C_K .
3. Iterate step 2 until no new elements are found.
4. Update $I_{uv} = I_{uv} \setminus C_K$.

Return K and $\{C_k\}_{k=1}^K$.

The iterative decreasing of the index set I_{uv} nicely illustrates that DBSCAN considers each element only once. The numerical most complex part of the algorithm is the calculation of the ε -neighbourhood. A naive approach would result in a complexity of $O(S_N^2)$ but this can be reduced to an average runtime of $O(S_N \log S_N)$ by smart implementations [109]. In [110] one uses the triangle inequality

$$\text{dist}(p, q) > |\text{dist}(p, r) - \text{dist}(q, r)|$$

for a distance function dist and points p, q, r to estimate a lower bound for the distance $\text{dist}(p, q)$ by using a reference point r . In this way, the number of distance calculations can be drastically reduced leading to a strong decrease in the runtime [110].

Figure 7.1 illustrates Algorithm 7.2.1 for $t_j \in \{0, \Delta_t, 2\Delta_t, \dots\}$ and $h_j \in \{0, \Delta_h, 2\Delta_h, \dots\}$ lying on a grid with step-size Δ_t and Δ_h . We simulated one probe, emitter and receiver in one, with positions given by $x_E(h) = x_R(h) = h$. We choose $\gamma = (\frac{\Delta_h}{\Delta_t})^2$ and $\varepsilon = \sqrt{2}\Delta_h$.

Given $\mathbf{t}_{j_1} = q_1\Delta_t$, $\mathbf{t}_{j_2} = q_2\Delta_t$, $h_{j_1} = p_1\Delta_h$ and $h_{j_2} = p_2\Delta_h$ with $p_1, p_2, q_1, q_2 \in \mathbb{N}$ we obtain

$$\begin{aligned} d_\gamma(\mathbf{t}_{j_1}, h_{j_1}, \mathbf{t}_{j_2}, h_{j_2}) &= \sqrt{\gamma(\mathbf{t}_{j_1} - \mathbf{t}_{j_2})^2 + \frac{1}{2}((x_E(h_{j_1}) - x_E(h_{j_2}))^2 + (x_R(h_{j_1}) - x_R(h_{j_2}))^2)} \\ &= \sqrt{\frac{\Delta_h^2}{\Delta_t^2}(q_1 - q_2)^2\Delta_t^2 + (p_1 - p_2)^2\Delta_h^2} \\ &= \Delta_h\sqrt{(q_1 - q_2)^2 + (p_1 - p_2)^2} \end{aligned}$$

and thus $(\mathbf{t}_{j_1}, h_{j_1})$ and $(\mathbf{t}_{j_2}, h_{j_2})$ are related, i.e. $d_\gamma(\mathbf{t}_{j_1}, h_{j_1}, \mathbf{t}_{j_2}, h_{j_2}) \leq \sqrt{2}\Delta_h$, if and only if $(\mathbf{t}_{j_2}, h_{j_2})$ is in the 8-neighbourhood of $(\mathbf{t}_{j_1}, h_{j_1})$.

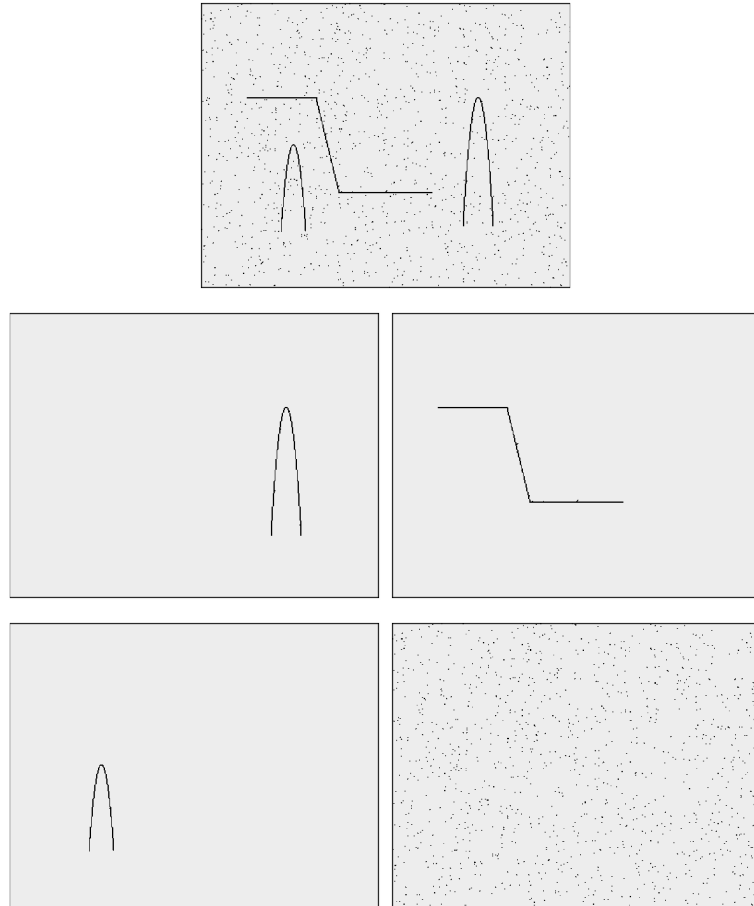


Fig. 7.1: top: original dataset, middle/bottom: solution obtained with DBSCAN: 3 clusters and noise.

7.3 Clustering Algorithms for Intersecting Defect Peaks

In the last section we presented an algorithm to determine the defect clusters for separated peaks from different defects. Let us consider the more general case where the peaks of different defects might intersect, i.e. for $k \neq k'$ there might be a point (t_j, h_j) such that

$$\int_{\partial D_k \cap E_{h_j, t_j}} a(t_j, y, h_j) dy \neq 0 \quad \text{and} \quad \int_{\partial D_{k'} \cap E_{h_j, t_j}} a(t_j, y, h_j) dy \neq 0$$

holds. This will change the requirements to our clustering algorithm. Here it is crucial for the reconstruction to use clustering methods that are able to handle overlapping clusters. Moreover, as peaks of different defects will now intersect, the relation measure (7.2) is no longer sufficient to separate unrelated elements. Thus, we need a clustering algorithm that can separate intersecting clusters. Furthermore, some points belong to several clusters and hence the algorithm should be able to handle multiple assignments. DBSCAN already does this somewhat at least for the low density points. Another approach is shown in [111] where the fuzzy-C-means algorithm is presented, a generalisation of k-means that uses fuzzy sets [112] where the membership of a point to a cluster can not only take the values 0 (does not belong to the cluster) or 1 (does belong to the cluster) but also values in between. However, such algorithms are useful to separate clusters that blend into each other, but they will not work when clusters e.g. cross each other as it might be the case in our application (see Figure 7.2). Here we need to modify the existing algorithms to make special use of the structure of the clusters. Therefore, we will only consider the ToFD application as overlapping signals do usually not occur in wall thickness measurements. In [113] a modified version of DBSCAN is presented that uses multiple density levels to decompose a set of points in several overlapping clusters. Thereby one exploits the fact that the clusters may have different densities.

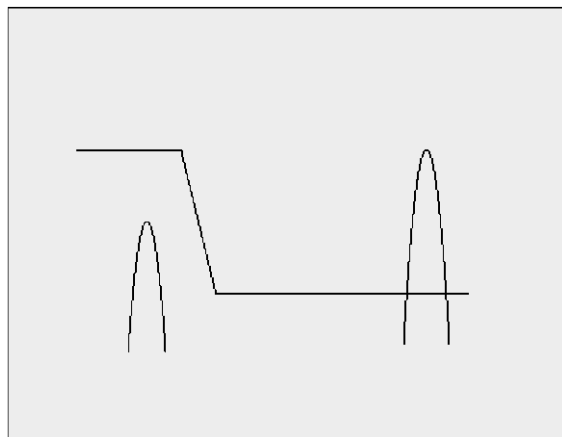


Fig. 7.2: Three overlapping clusters.

We will exploit the structure of our data by using the fact that our clusters are mostly shaped like a discretized curve in \mathbb{R}^2 . Therefore, we first rewrite DBSCAN in a different manner and then modify the algorithm with respect to our cluster structure. Remember that DBSCAN works as follows:

- Iterate through all points in the data.
- If a core-point is found, start a new cluster with the core-point and its ε -neighbourhood.
- Add the ε -neighbourhood of all core-points in the cluster to the cluster.
- Repeat this until the cluster does not change any longer. Then continue iterating through the data.

We note that the core-points form the basis of this algorithm. Because of that, we can rewrite the DBSCAN algorithm as

- Calculate the set C of all core-points in the data.
- Starting with a point in the set C as cluster, add all points in the ε -neighbourhood of the cluster that are also in C to the cluster.
- Repeat this until the cluster does not change any longer. Then start a new cluster with another core-point that isn't already assigned.
- For all found clusters, add the ε -neighbourhood to the cluster.

Note that the second and third step describe the Single-Linkage algorithm [104] applied to the set C of all core-points. Thus, DBSCAN first calculates all core points, then applies Single-Linkage and in the end adds all non-core-points to the clusters. In order to be able to separate intersecting clusters we will add another step that manipulates the set C in a special way before the Single-Linkage algorithm is applied. In the ToFD application, the clusters are mostly formed like curves in \mathbb{R}^2 with one dimension in time and one in the position of the probes. It follows that for the right choice of ε , a point in a cluster should typically have two other core-points, one that is previous and one that is successive on the curve. If the ε -neighbourhood contains less core-points, the point might be at one end of the cluster. Points that have more core-points in their ε -neighbourhood may indicate an intersection between two curves. Furthermore, we assume that the clusters are smooth in the sense that for two intersecting clusters with points P_1, Q, R_1 and P_2, Q, R_2 the angles $\sphericalangle(P_1QR_1)$, respectively $\sphericalangle(P_2QR_2)$, are closer to 180° than $\sphericalangle(P_1QR_2)$, respectively $\sphericalangle(P_2QR_1)$. In other words, coming to an intersection point of two clusters Q with an ε -neighbourhood containing more than 2 points, it is most likely that points belonging to the same cluster locally lie on a line-like curve. Using these information about our clusters we can perform the following manipulation of the set C . As long as there is a core-point P in C that has more than two neighbours corresponding to C :

- Find the two neighbours N_1, N_2 such that $\sphericalangle(N_1PN_2)$ is closest to 180° .

- Remove N_1, N_2 from the neighbourhood of P .
- Add a new point P' to C with neighbours N_1, N_2 .

Considering a point P with 5 neighbours, this iteration will replace P by three points P_1, P_2, P_3 that have at most two neighbours such that the angle formed by P_1, P_2, P_3 and their neighbours are near to 180° . Figure 7.3 illustrates the result. Note that one has to store the original point for each new point that was added to C as we do need to map each point back at the end.

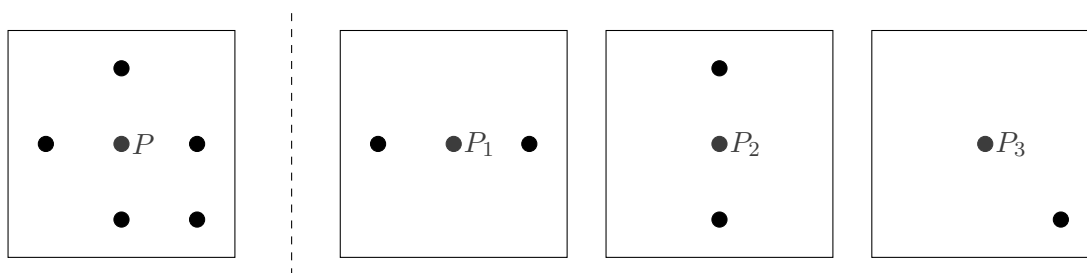


Fig. 7.3: left: P and its original ε -neighbourhood, right: copies P_1, P_2, P_3 of P and their neighbourhood.

Applying the Single-Linkage to the manipulated set C it will no longer merge intersecting clusters. Consider for example a horizontal and a vertical line that intersect at point P corresponding to Figure 7.3. While normally the complete neighbourhood of P would be added to a cluster and thus both lines would be assigned to the same cluster, with the new set of points P_1, P_2, P_3 , a line "coming from the right" will add P_1 to its cluster and continue on the left side while a line "coming from below" will add P_2 to its cluster and continue above.

However, after applying the Single-Linkage algorithm, we need to edit the returned clusters to adjust them to the original set of core points. First of all, for every cluster that contains points that were newly added to the set C , we remove these points and add the original point instead. As an example, for a cluster containing the point P_1 we remove the point P_1 and add P instead. But moreover, we have to delete some of the found clusters. Due to the manipulation of the set C the algorithm will return a lot of small clusters that are completely covered by other clusters. Considering Figure 7.3 again, we assume that P_1 and P_2 are assigned to two different clusters forming a horizontal and a vertical line. Furthermore, the 5th neighbour of P that was assigned to P_3 should be noise and thus P_3 and its neighbour will be returned as an independent cluster. After each point has been mapped back to P we obtain three clusters, one horizontal line containing P , one vertical line containing P and one cluster containing only P and one neighbour. It is clear that the last cluster is only a product of our manipulation of the set C and should be removed.

Altogether, using the above considered adaptations of DBSCAN we receive the following algorithm that we want to call DBCLAN as it is now more specialised in clustering lines and curves instead of only taking the spatial information into account.

Algorithm 7.3.1 (DBCLAN for intersecting defect reconstruction):

Given the pairs $\{(t_j, h_j)\}_{j=1}^{S_N}$, choose $\gamma > 0$, $\varepsilon > 0$, N_{del} arbitrarily and perform the following algorithm:

Determine the set C of all core-points, i.e. all points with at least 2 points in their ε -neighbourhood corresponding to the distance (7.2).

Manipulate the set C as long as there is a core-point with more then 2 core-points in its neighbourhood by

1. Select a point P in C with more then 2 core-points as neighbours.
2. Find the core-points N_1, N_2 in the neighbourhood of P with $\sphericalangle(N_1PN_2)$ closest to 180° .
3. Remove N_1 and N_2 from the neighbourhood of P .
4. Add a new point P' to C with neighbours N_1 and N_2 .

Apply Single-Linkage to C .

Map back all newly added points in C to their original point.

Add the ε -neighbourhood of each point in a cluster to the cluster. Here we also consider non-core-points.

Remove clusters that are completely covered by other clusters or clusters that have not more then N_{del} elements.

Return the set of remaining clusters and the number of clusters K .

The numerical most complex part of the algorithm is still the calculation of the ε -neighbourhood and thus the complexity stays at $O(S_N^2)$ respectively $O(S_N \log S_N)$ depending on the implementation. Figure 7.4 illustrates the algorithm for three clusters where two clusters intersect. We set $N_{\text{del}} = 5$ while the other parameters were chosen in the same manner as in the last section such that the ε -neighbourhood is equivalent to the 8-neighbourhood.

7.4 Summary

In this chapter we discussed algorithms to solve our second inversion step. Therefore a short introduction to the ideas of cluster analysis was given. We introduced the DBSCAN algorithm that can be seen as a generalization of Single-Linkage. This algorithm has been proved to be very useful for our application as it is not only able to reconstruct both the clusters and the number of clusters, but it also handles noisy data and thus has a denoising effect on our data. We adapted the algorithm to our application first with non-overlapping defects, see Algorithm 7.2.1. Later in this chapter we also considered intersecting defects. Taking the special form of our clusters into account we altered DBSCAN into DBCLAN 7.3.1. Besides the advantages of DBSCAN, this algorithm is

furthermore able to separate intersecting clusters. For both algorithms, 7.2.1 and 7.3.1, the numerical most expensive part is the calculation of the ε -neighbourhood. Using smart implementations this can be done in an average runtime of $O(S_N \log S_N)$, where S_N is the number of points obtained by the first reconstruction step. Because the number of points $S_N = O(N)$ only depends on the number of a-scans, this will allow a fast implementation of the algorithm.

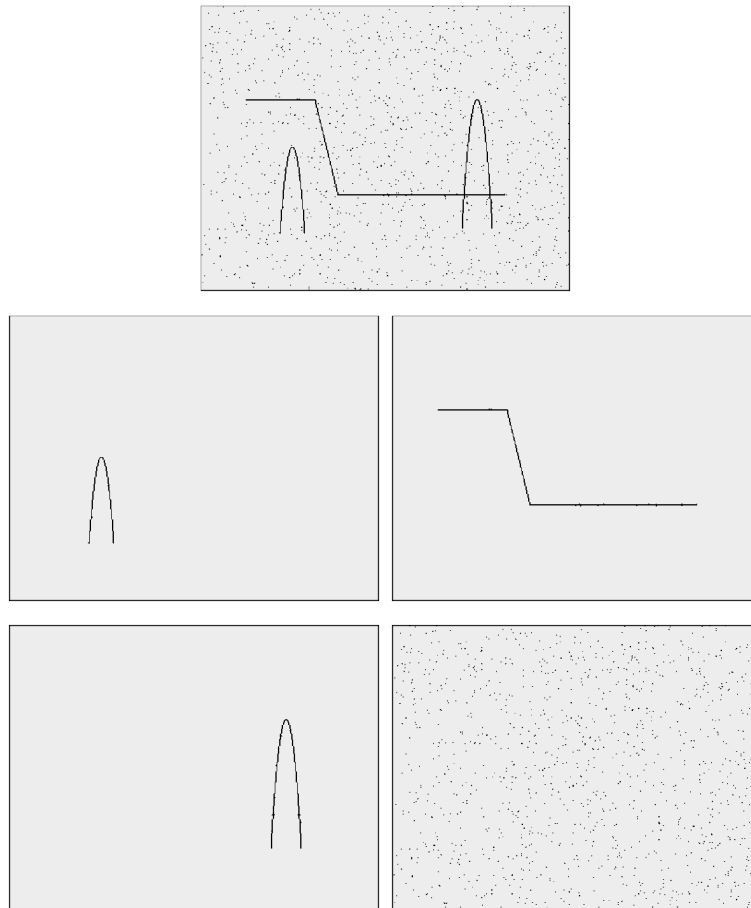


Fig. 7.4: top: original dataset, middle/bottom: solution obtained with DBLCAN: 3 clusters and noise.

8 Inversion 3 - Defect Reconstruction

In the last step of our inversion method we will finally use the data obtained in step 2 to reconstruct the defect boundary. Given the result of the second inversion step we have several clusters, each representing one defect. Thus we only need a reconstruction method for a single cluster. This method can then be applied to each cluster separately. In this way, we do not only obtain a simpler problem with only one defect in step 3 but this also allows a parallel computing of several defects and hence a fast implementation.

Let us consider one cluster obtained in the second inversion, i.e. a set of points $\{\mathbf{t}_k, h_k\}_{k=1}^N$ where N is the number of points in the cluster. We denote the defect belonging to this cluster by D and omit the defect indices as the inversion method will reconstruct each defect on its own. Thus, the problem can be stated as follows. Given the points $\{\mathbf{t}_k, h_k\}_{k=1}^N$ with

$$\int_{\partial D \cap E_{h_k, \mathbf{t}_k}} a(\mathbf{t}_k, y, h_k) dy \neq 0 \quad (8.1)$$

we want to reconstruct ∂D . We note that the function a derived in Chapter 4 has a complicated structure that highly depends on ∂D itself. To overcome this difficulty we will not use the information (8.1) directly but use that

$$\partial D \cap E_{h_k, \mathbf{t}_k} \neq \emptyset \quad (8.2)$$

follows from (8.1) for all $k = 1, \dots, N$. Furthermore, we need a discrete representation of ∂D . Let us therefore introduce the following definitions.

Definition 8.0.1:

Remember that in this work $\mathbb{R}_+ = [0, \infty)$. Let $D \subset \mathbb{R}^n$ be a bounded, connected set that represents a defect in an n -dimensional model space with $n = 2, 3$. We call D *star-shaped* if there exists a center point $C \in D$ and a continuous function $r : \mathcal{S} \rightarrow \mathbb{R}_+$ on the n -dimensional sphere \mathcal{S} such that

$$\partial D = \{C + r(\zeta)\zeta \mid \zeta \in \mathcal{S}\}.$$

We say that D is *non-degenerate* if r is Lipschitz-continuous and strictly positive, i.e. $C \notin \partial D$. Furthermore, let $\{\zeta_j\}_{j=1}^M$ be a discrete sampling of \mathcal{S} . We call the set

$$\{P_j = C + r_j\zeta_j\}_{j=1}^M \subset \partial D$$

with $r_j = r(\zeta_j)$ a *discrete representation* of ∂D . Let $\Lambda(\zeta_1, \dots, \zeta_M)$ be a polygonal line ($n = 2$) or triangulation ($n = 3$) of \mathcal{S} using the points $\{\zeta_j\}_{j=1}^M$. We say that $\Lambda(P_1, \dots, P_M)$ is an *approximation of ∂D* where $\Lambda(P_1, \dots, P_M)$ is defined by

$$\begin{aligned} \overline{P_i P_j} \in \Lambda(P_1, \dots, P_M) &\Leftrightarrow \overline{\zeta_i \zeta_j} \in \Lambda(\zeta_1, \dots, \zeta_M) & (n = 2), \\ \Delta(P_i, P_j, P_k) \in \Lambda(P_1, \dots, P_M) &\Leftrightarrow \Delta(\zeta_i, \zeta_j, \zeta_k) \in \Lambda(\zeta_1, \dots, \zeta_M) & (n = 3). \end{aligned}$$

We call $\Lambda(P_1, \dots, P_M)$ *suitable* if it is also star-shaped and non-degenerate. This does not necessarily hold as e.g. for $n = 2$ and $M = 2$ the set $\Lambda(P_1, \dots, P_M) = \Lambda(P_1, P_2)$ degenerates into a line.

Note that $\Lambda(P_1, \dots, P_M)$ is again a polygonal line / triangulation of ∂D but using the definition above it can be derived without knowledge about ∂D . This will be useful for our reconstruction method since we do not know ∂D . A 2-dimensional example of a defect D , a discrete representation and its polygonal line is given in Figure 8.1. Furthermore, we can prove the following theorem.

Theorem 8.0.2:

Let D be a non-degenerate, star-shaped defect and let $\{\zeta_j\}_{j=1}^M$ be a discrete sampling of \mathcal{S} such that

$$\min_j \|\zeta_j - \zeta\|_2 \leq \varepsilon$$

holds for an $\varepsilon > 0$ and all $\zeta \in \mathcal{S}$. Let $\{P_j = C + r_j\zeta_j\}_{j=1}^M$ be a representation of ∂D . Then

$$\max_{P \in \partial D} \min_j \|P - P_j\|_2 \leq (R + L)\varepsilon$$

holds with $R = \max_{\zeta \in \mathcal{S}} r(\zeta)$ and the Lipschitz constant L of the radius function r .

Proof:

Choose $P = C + r(\zeta)\zeta \in \partial D$ arbitrary. It follows

$$\begin{aligned}\|P - P_j\|_2 &= \|C + r(\zeta)\zeta - C - r_j\zeta_j\|_2 \\ &\leq \|r(\zeta)\zeta - r_j\zeta\|_2 + \|r_j\zeta - r_j\zeta_j\|_2 \\ &= |r(\zeta) - r(\zeta_j)| + r_j\|\zeta - \zeta_j\|_2 \\ &\leq L\|\zeta - \zeta_j\|_2 + R\|\zeta - \zeta_j\|_2\end{aligned}$$

and thus

$$\min_j \|P - P_j\|_2 \leq (L + R) \min_j \|\zeta - \zeta_j\|_2 \leq (L + R)\varepsilon.$$

□

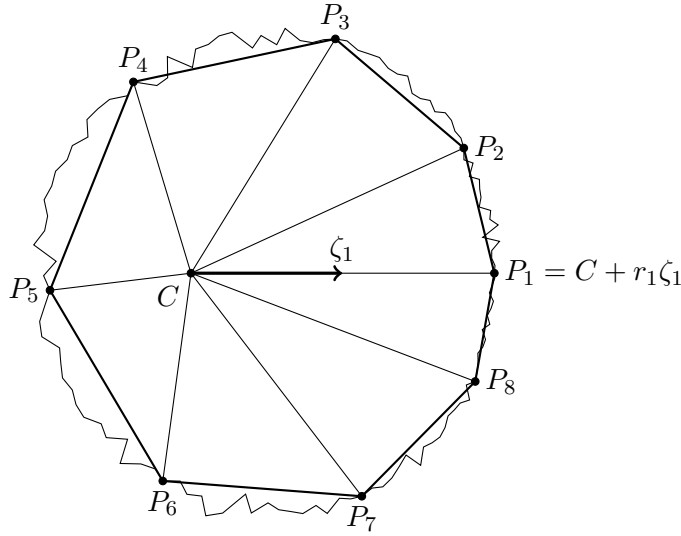


Fig. 8.1: Defect boundary $\partial D \subset \mathbb{R}^2$, a discrete representation with points $\{P_j\}_{j=1}^8$ and its polygonal line.

At this point we like to clarify the at first sight slightly conflicting notation. Theorem 8.0.2 shows an approximation property of the representation $\{P_j\}_{j=1}^M$ of ∂D . This might be confusing as there is also the set $\Lambda(P_1, \dots, P_M)$ which we call the approximation of ∂D . The reason for this notation is simple: $\{P_j\}_{j=1}^M$ is a discrete subset of ∂D and thus a representation of ∂D in the sense of Theorem 8.0.2. However, if one wants to analyse special characteristics of the defect, like area, volume or diameter, the representation $\{P_j\}_{j=1}^M$ is no longer sufficient. Thus we will use the set $\Lambda(P_1, \dots, P_M)$ as an approximation to recover those characteristics. Furthermore, we like to note that in Theorem 8.0.2 the set $\{P_j\}_{j=1}^M$ can be replaced by $\Lambda(P_1, \dots, P_M)$ and the assertion still holds since $\{P_j\}_{j=1}^M \subset \Lambda(P_1, \dots, P_M)$.

Let us now continue our considerations. Given Definition 8.0.1 we can try to reconstruct ∂D by finding a representation $\{P_j\}_{j=1}^M$ for which (8.2) holds. We obtain the discrete problem, given $\{t_k, h_k\}_{k=1}^N$ with

$$\{P_j\}_{j=1}^M \cap E_{h_k, t_k} \neq \emptyset, \quad (8.3)$$

find C , $\{r_j\}_{j=1}^M$ and $\{\zeta_j\}_{j=1}^M$ where M is also unknown. Note that the condition (8.3) can be reformulated as

$$\forall k \leq N : \exists j \leq M : P_j \in E_{h_k, t_k}. \quad (8.4)$$

After the points $\{P_j\}_{j=1}^M$ have been reconstructed, we can use $\Lambda(P_1, \dots, P_M)$ as an approximation for ∂D to extract the important information about D .

The algorithm to reconstruct the points $\{P_j\}_{j=1}^M$ presented in this chapter is divided in two parts. In a first step, we approximate the center point C as this is the reference point for all points P_j , $j = 1, \dots, M$. This can be done, using heuristics that are highly adapted to the measurement set-up. In the first section of this chapter we will present three methods to calculate C starting with one that is adapted to ToFD. The second method is adapted to wall thickness measurements and the third method is applicable in a more general set-up. If an approximation of C is derived, we can reconstruct $\{r_j\}_{j=1}^M$ and $\{\zeta_j\}_{j=1}^M$ in a second step. Using some a-priori conditions for the defect D the reconstruction will be reformulated as a minimization problem of special form which can be solved with an iterative method. This algorithm will be presented in Section 8.2.

8.1 Reconstruction of the Center Point

The point C plays an important role in the representation of ∂D . Thus, before we can calculate the points P_j , we first need to approximate C . Thereby the interesting question occurs, how much an approximation error in C influences the accuracy of the second part, the reconstruction of $\{r_j\}_{j=1}^M$ and $\{\zeta_j\}_{j=1}^M$. Hence let us state the following theorem.

Theorem 8.1.1:

Let $D \subset \mathbb{R}^n$ be a non-degenerate, star-shaped defect. For a suitable approximation $\Lambda(P_1, \dots, P_M)$ of ∂D define the set \mathfrak{C} of all center points of $\Lambda(P_1, \dots, P_M)$ for which a continuous, strictly positive function r exists, such that

$$\Lambda(P_1, \dots, P_M) = \{C + r(\zeta)\zeta \mid \zeta \in \mathcal{S}\}$$

holds. Then $\mathfrak{C} \neq \emptyset$ is open.

Proof:

For suitable approximations of ∂D the set \mathfrak{C} is non-empty by definition. We choose

$C \in \mathfrak{C}$. Note that $C \in \mathfrak{C}$ if and only if for all $P \in \Lambda(P_1, \dots, P_M)$ the line $\overline{CP} \setminus \{P\}$ is in the interior of $\Lambda(P_1, \dots, P_M)$ since it follows directly that for each direction $\zeta \in \mathcal{S}$ there is exactly one point $P \in \Lambda(P_1, \dots, P_M)$ that can be written as $P = C + r\zeta$ with $r > 0$. Hence the radius function r is defined point-wise. The continuity of r follows from the continuity of $\Lambda(P_1, \dots, P_M)$ and $r > 0$ holds since C is in the interior of $\Lambda(P_1, \dots, P_M)$.

We first consider the two-dimensional case $n = 2$. For fixed k, l , let E' be the affine hyperplane that contains the line $\overline{P_k P_l} \in \Lambda(P_1, \dots, P_M)$. E' divides \mathbb{R}^2 in two half spaces, denote the half space that contains C by E excluding E' . Note that E is open. Now let $\{E_j\}_{j=1}^L$ be the set of all half spaces that can be generated in this way for all k, l such that $\overline{P_k P_l} \in \Lambda(P_1, \dots, P_M)$ and $\{E'_j\}_{j=1}^L$ the corresponding affine hyperplanes. We show that $\mathfrak{C} = \cap_{j=1}^L E_j$ and thus open.

Therefore, let $C' \notin \cap_{j=1}^L E_j$, this means $C' \notin E_{j'}$ for at least one $j' \leq L$. Let $\overline{P_k P_l}$ be the line that is contained in $E'_{j'}$ and choose $P \in \overline{P_k P_l}$ with $P \neq P_k$ and $P \neq P_l$. Then there exist an $\varepsilon > 0$ such that $B_\varepsilon(P) \cap E_{j'}$ is in the interior and $B_\varepsilon(P) \setminus E_{j'}$ is not in the interior of $\Lambda(P_1, \dots, P_M)$. Since $C' \notin E_{j'}$ and $E_{j'}$ is a half plane, it follows that $\overline{C'P} \cap (B_\varepsilon(P) \setminus E_{j'}) \neq \{P\}$, i.e. $\overline{C'P}$ contains points unequal to P that are not in the interior of $\Lambda(P_1, \dots, P_M)$. Hence $C' \notin \mathfrak{C}$.

Now choose $C' \notin \mathfrak{C}$. It follows that there exist points $P \in \partial D$ and $C_1 \in \overline{C'P}$ such that $\overline{C_1 P} \subset \overline{C'P}$ is not in the interior of D . As above we obtain that $\overline{C'P} \cap (B_\varepsilon(P) \setminus E_{j'}) \neq \{P\}$ following that $C' \notin E_{j'}$. Thus we obtain $\mathfrak{C} = \cap_{j=1}^L E_j$.

For $n = 3$ this follows with the same argumentation but using for $j, k, l \leq M$ the affine hyperplanes that contain the triangles $\Delta(P_j, P_k, P_l) \in \Lambda(P_1, \dots, P_M)$. \square

We obtain

Corollary 8.1.2:

For a given defect $D \subset \mathbb{R}^n$ let $\Lambda(P_1, \dots, P_M)$ be a suitable approximation. Furthermore, let \tilde{C} be an approximation of a center point $C \in \mathfrak{C}$ of $\Lambda(P_1, \dots, P_M)$ and define the approximation error

$$\varepsilon := \|C - \tilde{C}\|_2.$$

Then $\tilde{C} \in \mathfrak{C}$ if ε is small enough.

Proof:

This follows directly from the fact that \mathfrak{C} is open. \square

Note that we cannot provide any bound for the approximation error in the corollary. Indeed, the size of the set \mathfrak{C} strongly depends on the shape of the defect. Figure 8.2 shows an example where \mathfrak{C} becomes small. However, the defects appearing in practice are mostly “well-shaped” and thus the set \mathfrak{C} is relatively big. Hence, referring to Corollary 8.1.2, we will from now on assume that C is given exactly.

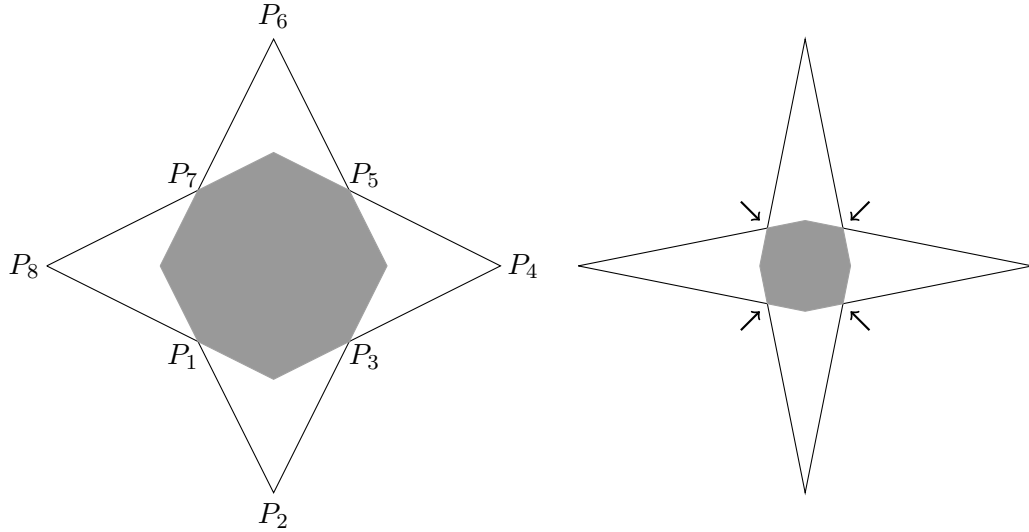


Fig. 8.2: A star shaped defect and the set \mathfrak{C} (gray). The set becomes smaller if the points P_1, P_3, P_5, P_7 approach each other.

After these theoretical considerations about the behaviour of the set \mathfrak{C} , we are going to discuss several methods in the following subsections. Thereby, the diversity of defects and applications makes it necessary to exploit given a-priori knowledge and use heuristics to develop approximation algorithms for C . As there is no way to prove if C is a center point or not, all methods presented here try to find a point that is somewhere in the middle of the defect. For defect shapes that appear in practical applications, this is a good approximation to a center point. Although these methods have no approximation guarantee, the results obtained in numerical examples are quite good.

8.1.1 A Heuristic Method for ToFD

Let us start considering a heuristic method for the ToFD technique. Therefore remember that the position of emitter and receiver for a pair (t, h) is given by the (known) functions $x_E(h)$ and $x_R(h)$. For a set of points $\{t_k, h_k\}_{k=1}^N$ let us define the set

$$\mathfrak{r} := \{(x_E(h_k), x_R(h_k)) \mid t_k = \min_i t_i\}$$

of emitter and receiver positions where the measuring time t_k is minimal. As the measuring time in homogeneous materials directly correlates with the distance between emitter, reflection/diffraction point and receiver, the set \mathfrak{r} contains all positions where the probes were closest to the defect. We can now use the fact that ToFD is mainly used to analyse welds in tubes. Thereby, the probes are aligned geometrically on each side of the weld. It follows that for $(x_E, x_R) \in \mathfrak{r}$ the middle point $x_M := \frac{x_E + x_R}{2}$ lies in the “weld-plane”. Assuming that defects may only appear in the weld but not in the tube itself, the point

x_M is directly above the reflection/diffraction point P . Knowing that for a pair (t, h) the equation $\|x_E - P\|_2 + \|x_R - P\|_2 = ct$ holds and using the symmetry of the arrangement we obtain by Pythagoras $P = x_M + \frac{1}{2}\sqrt{(ct)^2 - \|x_E - x_R\|_2^2}\zeta_M$ where ζ_M is the direction “downward the weld-plane” that is perpendicular to the probes movement and to $x_R - x_E$. Thus we can calculate one point $P \in \partial D$ on the defect boundary. Figure 8.3 illustrates this.

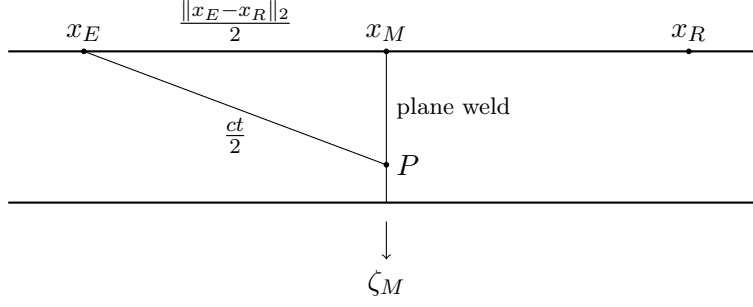


Fig. 8.3: Calculation of the point $P \in \partial D$ for $(x_E, x_R) \in \mathfrak{r}$.

Next we have to decide whether the constructed point P is on the lower or on the upper part of the defect, i.e. if we have measured a diffraction or reflection. Remember that due to the amplitude function \mathcal{A} a reflection will most likely have a much smaller amplitude than a diffraction but a reflection from the surface of a defect will arrive before the diffraction from its bottom. For the probe positions $(x_E(h_j), x_R(h_j)) \in \mathfrak{r}$ consider the set $T := \{t_k \mid h_k = h_j\}$ of all arrival times that were measured in the same a-scan. Now we have to analyse two cases:

1. There is no arrival time $t_k \in T$ that is significantly greater than t_j , i.e there is no $t_k \in T$ such that $t_k > t_j + \varepsilon$ holds. In this case, the only signals we have measured in this a-scan are coming from approximatively the same positions. As the diffraction has a stronger amplitude than the reflection, we assume that (t_j, h_j) was a diffraction signal and the reflection got lost in noise. Thus we choose $C = P - \delta\zeta_M$ as a point that is above the point P on the lower part of the defect boundary. Here δ has to be chosen appropriately.
2. In the second case there is an arrival time $t_k \in T$ for which $t_k > t_j + \varepsilon$ holds. Here we have measured signals coming from two or more different regions of the defect. Assuming that the first signal is now a reflection measure, we set $C = P + \delta\zeta_M$ below the point P that is on the upper part of the defect boundary.

Note that if the set \mathfrak{r} contains more than one pair of positions, we can use the method to calculate several center points C . As shown in Theorem 8.1.1, the set \mathfrak{C} of all center points is convex and hence we can derive the mean of all generated points C . This will enhance the algorithm and make it more stable against noise. We obtain

Algorithm 8.1.3 (Center point approximation for ToFD):

Given the pairs $\{t_k, h_k\}_{k=1}^N$ that have been derived by the first and second part of our reconstruction method applied to ToFD-measurement data, we perform the following steps to find an approximation C of a center point:

Choose $\varepsilon, \delta > 0$ appropriately.

Initialize $C = 0$ and $\zeta_M \in \mathcal{S}$ perpendicular to probe movement in the weld plane.

Calculate $\mathfrak{r} := \{(x_E(h_k), x_R(h_k)) \mid t_k = \min_i t_i\}$.

For all $(x_E(h_j), x_R(h_j)) \in \mathfrak{r}$:

Calculate $T := \{t_k \mid h_k = h_j\}$.

If $\exists t_k \in T : t_k - t_j > \varepsilon$

then $C = C + \frac{x_E(h_j) + x_R(h_j)}{2} + \frac{1}{2} \sqrt{(ct_j)^2 - \|x_E(h_j) - x_R(h_j)\|_2^2} \zeta_M + \delta \zeta_M$

else $C = C + \frac{x_E(h_j) + x_R(h_j)}{2} + \frac{1}{2} \sqrt{(ct_j)^2 - \|x_E(h_j) - x_R(h_j)\|_2^2} \zeta_M - \delta \zeta_M$.

Average $C = \frac{C}{|\mathfrak{r}|}$.

8.1.2 A Heuristic Method for Wall Thickness Measurement

As a second example for highly adapted heuristic methods we want to consider a technique to reconstruct C for wall thickness measurement data. Therefore remember that this ultrasonic non-destructive testing technique is a classical application of an impulse echo method. This means a probe, emitter and receiver in one, placed at position x_E sends out the ultrasound in direction ζ_M of the tube back wall. Thus, the measured signals are most likely reflections from points below the tube in direction ζ_M or near by. Given the set $\{t_k, h_k\}_{k=1}^N$ we define the mean probe position $x_M = \frac{1}{N} \sum_{k=1}^N x_E(h_k)$. This is the mean of all positions where defect reflections were measured. Referring to the considerations above, the position x_M is directly atop the middle of the defect. Neglecting the curvature of the material we obtain that x_M is on the surface and hence the alternate point on the back wall is given by $C = x_M + s_t \zeta_M$ where s_t is the thickness of the component. C will surely be inside the defect as it is on the back wall and we are analysing back wall defects and so we can use it as an approximation for a center point. We obtain the method

Algorithm 8.1.4 (Center point approximation for back wall defects):

Given the pairs $\{t_k, h_k\}_{k=1}^N$ that were derived by the first and second part of our reconstruction method applied to wall thickness measurement data, we use

$$C = \frac{1}{N} \sum_{k=1}^N x_E(h_k) + s_t \zeta_M$$

as an approximated center point. Here s_t is the diameter of the component and ζ_M is the directional vector pointing from surface to back wall.

8.1.3 A General Heuristic Approach

The last reconstruction method for C we want to present is a general heuristic approach that can be applied to almost every sort of testing set-up. Therefore, we use that for most defect shapes appearing in practical applications the “middle” of the defect is a good guess for a center point $C \in \mathfrak{C}$. Let us define the middle of a defect D as those points which are in the interior of D and have the greatest distance to the boundary ∂D , i.e.

$$\begin{aligned} \mathfrak{M} &:= \{C \in D \mid \text{dist}(C, \partial D) \geq \text{dist}(C', \partial D) \text{ for all } C' \in D\} \\ &= \{C \in D \mid \min_{P \in \partial D} \|C - P\|_2 \geq \min_{P \in \partial D} \|C' - P\|_2 \text{ for all } C' \in D\}. \end{aligned}$$

We assume $\mathfrak{M} \subset \mathfrak{C}$ and approximate a center point $C \in \mathfrak{M}$. Because ∂D is unknown, we cannot determine $\text{dist}(C, \partial D)$. Thus we use the knowledge that for each pair (t_k, h_k) there exists a point $P \in E_{t_k, h_k} \cap \partial D$ where E_{t_k, h_k} is given by Definition 4.1.1. For a given set of points $\{t_k, h_k\}_{k=1}^N$ we can now approximate

$$\text{dist}(C, \partial D) \approx \min_k \text{dist}(C, E_{t_k, h_k}) =: R(C)$$

with $R : \mathbb{R}^n \rightarrow \mathbb{R}_+$. We seek for a point $C \in \mathfrak{M}$ where the distance to the boundary ∂D is maximal, hence we want to find a local maximum of $R(C)$. Note that for $\|C\|_2 \rightarrow \infty$ also $R(C) \rightarrow \infty$ and thus R has no global maximum. To avoid this problem, we multiply R with an indicator function $I : \mathbb{R}^n \rightarrow [0, 1]$ where $I(C)$ is a probability function that has to be defined a-priori. It should be chosen in a way such that $I(C)$ is great for $C \in D$ and $R(C) \cdot I(C)$ is bounded. Figure 8.4 illustrates R , I and $R \cdot I$ for the two dimensional area of size $10\text{mm} \times 10\text{mm}$. We have chosen the pairs $\{(t_k, h_k)\}_{k=1}^4 = \{(1.35\mu\text{s}, 0\text{mm}), (0.84\mu\text{s}, 5\text{mm}), (1.35\mu\text{s}, 10\text{mm}), (2.7\mu\text{s}, 5\text{mm})\}$ and $x_E(h_k) = x_R(h_k) = (h_k, 0\text{mm})$. For the left example we neglected the last point $(2.7\mu\text{s}, 5\text{mm})$. We observe that the function $R(C)$ in both cases is unbounded for $\|C\|_2 \rightarrow \infty$ while $R(C) \cdot I(C)$ has a global maximum where we chose I as an exponential decreasing function that has its maximum in the center of the plane.

We now can derive $C \in \mathfrak{C}$ by

$$\begin{aligned} C &= \arg \max_{C' \in \mathbb{R}^n} R(C')I(C') \\ &= \arg \max_{C' \in \mathbb{R}^n} \min_k \text{dist}(C', E_{t_k, h_k})I(C') \\ &= \arg \min_{C' \in \mathbb{R}^n} \max_k -\text{dist}(C', E_{t_k, h_k})I(C'). \end{aligned}$$

This is a so called *minimax problem*. To solve this problem one uses the fact that for continuous differentiable functions $\{f_k\}_{k=1}^N$ the directional derivatives of $q(x) := \max_k f_k(x)$ exists and can be determined. Thus solving the minimax problem means solving the problem $\min_x q(x)$ what can be done by applying an iterative method using the

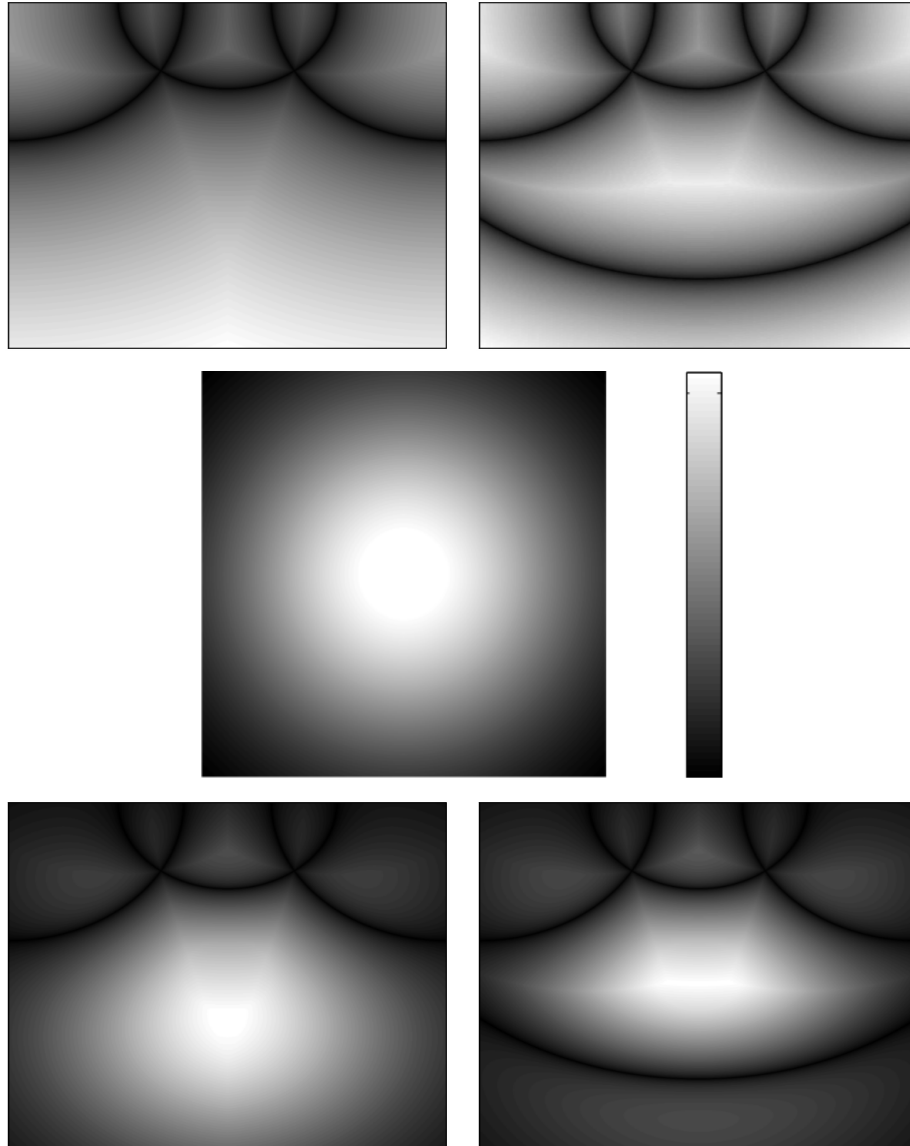


Fig. 8.4: Top: The function $R(C)$, $C \in \mathbb{R}^2$ for $N = 3$ (left) and $N = 4$ (right) given pairs $\{(t_k, h_k)\}_{k=1}^N$. Middle: The function $I(C)$. Bottom: $R(C) \cdot I(C)$.

direction of steepest descend that can be derived with help of the directional derivatives. For more information about minimax problems, its theory and numerical algorithms we refer to [114–116] and the references therein.

We have now introduced three different methods to approximate a point $C \in \mathfrak{C}$. One may also think of other heuristic ideas e.g. for other testing set-ups. However, we may stay with those three examples and change to the second part of reconstruction step 3, the reconstruction of the points $\{P_k\}_{k=1}^M$ with $P_k = C + r_k \zeta_k$ where r_k , ζ_k and M are unknown.

8.2 Spiderweb Algorithm for Defect Reconstruction

In the last section we discussed the reconstruction of $C \in \mathfrak{C}$ and due to Corollary 8.1.2 we assume in this section that C is exactly given. According to (8.4) we obtain the discrete problem: Given $C \in \mathfrak{C}$ and $\{\mathbf{t}_k, h_k\}_{k=1}^N$, find $\{r_j\}_{j=1}^M$ and $\{\zeta_j\}_{j=1}^M$ such that

$$\forall k \leq N : \exists j \leq M : P_j \in E_{\mathbf{t}_k, h_k} \quad (8.5)$$

where $P_j = C + r_j \zeta_j$. Note that we do only need one point P_j for each pair (\mathbf{t}_k, h_k) and hence our algorithm will reconstruct not more than $M \leq N$ points. Due to the characteristic visualization of its iteration (see Figure 8.5) we like to call our method the *spiderweb algorithm*. We will introduce the method in several steps. First we deduce a minimization problem from the above stated problem and analyse its solution. In a second step we show that for exactly given $\{\zeta_j\}_{j=1}^M$ the solution of the minimization problem correlates with the solution of (8.5). Furthermore we show that our algorithm is stable in the sense that given an approximation of $\{\zeta_j\}_{j=1}^M$ the reconstructed points will be an approximation to the solution of (8.5). We denote the scenario of (approximately) given $\{\zeta_j\}_{j=1}^M$ as *supervised minimization*. In a last step we discuss the case of *unsupervised minimization* where $\{\zeta_j\}_{j=1}^M$ is not given. We can prove that, taking a dense sampling of \mathcal{S} as input for our method, our algorithm will automatically reduce the number of directions and determine a subset that is an approximation to $\{\zeta_j\}_{j=1}^M$.

8.2.1 Reconstruction as Minimization Problem

One can easily see that the problem (8.5) has no unique solution. For a given set of points $\{P_j\}_{j=1}^M$ that satisfies (8.5) we can add a point P_{M+1} and obtain another solution $\{P_j\}_{j=1}^{M+1}$. Hence, we need to specify the solution we want to reconstruct. Let us therefore define

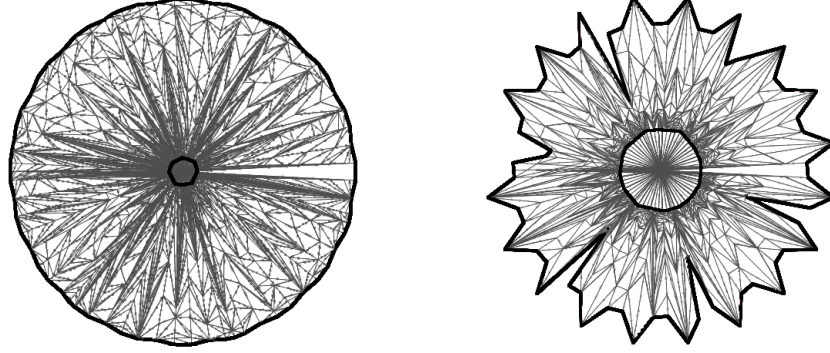


Fig. 8.5: Two visualizations of the iterations of our algorithm in \mathbb{R}^2 . The shape is formed like a spiderweb. Here marked: The starting polygon (thick outer line) and the returned solution (thick inner line).

Definition 8.2.1:

We call $\{P_j\}_{j=1}^M$ an *irreducible solution* for $\{(t_k, h_k)\}_{k=1}^N$ if

$$\forall k \leq N : \exists j \leq M : P_j \in E_{t_k, h_k} \quad (8.6)$$

but for all $j' \leq M$ the set $\{P_j\}_{j=1, j \neq j'}^M$ does not satisfy (8.6). Furthermore let \mathfrak{P} be the set of all irreducible solutions and $V : \mathfrak{P} \rightarrow \mathbb{R}$. We say that $\{P_j\}_{j=1}^M$ is minimal according to V if

$$V(\{P_j\}_{j=1}^M) = \min_{\mathfrak{P}} V.$$

Note that for every set $\{(t_k, h_k)\}_{k=1}^N$ the points $\{P_k\}_{k=1}^N$ with $P_k \in E_{t_k, h_k}$ generate a solution for (8.5). It follows directly that there always exists a irreducible solution since we can remove points from the set $\{P_k\}_{k=1}^N$ until the solution is irreducible. Now we assume that a defect normally has a small size, and we design the algorithm to reconstruct an irreducible solution that is minimal according to a function V that will be discussed more precisely later. For a better understanding the reader might think of V being the volume of the approximation $\Delta(P_1, \dots, P_M)$. We thus assume that the desired reconstruction is also a solution of the problem

$$\min_{\{P_j\}_{j=1}^M \in \mathfrak{P}} V(\{P_j\}_{j=1}^M). \quad (8.7)$$

We will use this minimization approach to construct the spiderweb algorithm that determines a set of radii $\{r_j\}_{j=1}^M$ for a given set of directions $\{\zeta_j\}_{j=1}^M$. Thus, assume the directions $\{\zeta_j\}_{j=1}^M$ are given and fixed. Then the point P_j lies on the ray $\{C + r\zeta_j \mid r > 0\}$. The condition $P_j \in E_{t_k, h_k}$, i.e. P_j is in the intersection of a line and a spheroid, becomes a non-linear equation with respect to r . We can define

Definition 8.2.2:

Let the set $\{t_k, h_k\}_{k=1}^N$, the center point C and the directions $\{\zeta_j\}_{j=1}^M$ be given. Furthermore let $B_{kj} = \{r \mid C + r\zeta_j \in E_{t_k, h_k}\}$ be the set of all solutions of the equation $P_j \in E_{t_k, h_k}$. We define

$$b_{kj} = \min((B_{kj} \cap \mathbb{R}_+) \cup \{\infty\})$$

as the smallest, real and positive solution in B_{kj} or ∞ if no positive real solution exists.

Note that for (t_k, h_k) being a diffraction signal, the set B_{kj} will contain positive and negative solutions as C is inside the ellipsoid E_{t_k, h_k} while it is outside for reflection signals. Thus the algorithm will automatically distinguish between both cases. Figure 8.6 illustrates the value b_{kj} for different cases in two dimensions. Moreover, if $x_E(h_k) = x_R(h_k)$, the set E_{t_k, h_k} is a sphere and thus B_{kj} is explicitly given by

$$B_{kj} = \left\{ \langle x_E(h_k) - C, \zeta_j \rangle \pm \sqrt{\langle x_E(h_k) - C, \zeta_j \rangle^2 - \|x_E(h_k) - C\|_2^2 + \frac{c^2 t_k^2}{4}} \right\}.$$

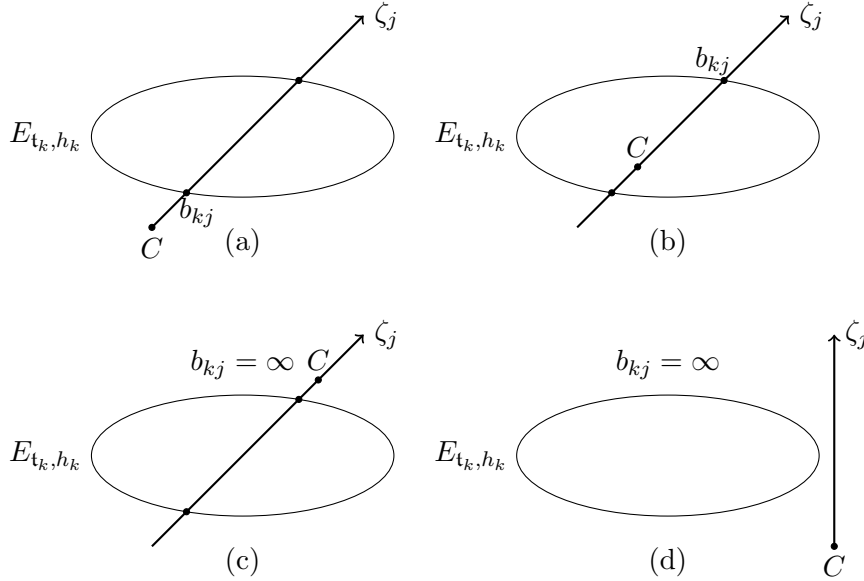


Fig. 8.6: The solution b_{ki} for different cases: (a) B_{kj} contains two positive solutions, b_{kj} is the smaller one; (b) B_{kj} contains one negative and one positive solution, b_{kj} is the positive solution; (c) B_{kj} contains two negative solutions, b_{kj} is set to ∞ ; (d) B_{kj} contains no (real) solution, b_{kj} is set to ∞ .

Furthermore we denote $r = (r_j)_{j=1}^M$ and $b_k = (b_{kj})_{j=1}^M$. As the directions $\{\zeta_j\}_{j=1}^M$ and the center point C are fixed, we can rewrite the function V as function in the radii,

i.e. $V : \mathbb{R}^M \rightarrow \mathbb{R}$ with $V(r) = V(\{P_j\}_{j=1}^M)$. Using Definition 8.2.2 we can rewrite minimization problem (8.7) as

$$\min\{V(r) \mid r \in \mathbb{R}^M, \forall k \leq N : \exists j \leq M : r_j \in B_{kj}\}. \quad (8.8)$$

Having in mind that V will favour small defects, i.e. defects with small radii r , we relax the constraints of the above stated problem and instead consider

$$\begin{aligned} & \min\{V(r) \mid r \in \mathbb{R}^M, \forall k \leq N : \exists j \leq M : r_j \geq b_{ki}\} \\ \Leftrightarrow & \min\{V(r) \mid r \in \mathbb{R}^M, \forall k \leq N : r \not\prec b_k\} \end{aligned} \quad (8.9)$$

where $r \not\prec b_k$ is meant in a component-by-component comparison sense. Note that for a solution r of (8.9) and a solution \tilde{r} of (8.8) $V(r) \leq V(\tilde{r})$ holds. We will show in the next subsection that the solutions of both problems are equal under special conditions. However, motivated by problem (8.9), we define

Definition 8.2.3:

For $k = 1, \dots, N$ let $Q_k := \{r \in \mathbb{R}^M \mid r < b_k\}$ be the set of all vectors that are smaller than b_k in a component-by-component sense. We define $Q = \cup_{k=1}^N Q_k \cup (\mathbb{R}^M \setminus \mathbb{R}_+^M)$ as the union of all Q_k and all vectors in \mathbb{R}^M with at least one negative component. We can rewrite (8.9) as

$$\min_{r \in \mathbb{R}^M \setminus Q} V(r). \quad (8.10)$$

Furthermore, for $j = 1, \dots, M$ let $e_j \in \mathbb{R}^M$ be the j -th canonic unit vector. We call $r \in \mathbb{R}^M$ a *concave corner of Q* if $r \notin Q$ but $r - \varepsilon e_j \in Q$ for all $\varepsilon > 0$ and all $j = 1, \dots, M$. Figure 8.7 illustrates this for $M = 2, N = 4$ (top) and $M = N = 3$ (bottom). Moreover for $j = 1, \dots, M$ we define

$$b_{kj} := \begin{cases} \infty & k = N + 1, \dots, N + M, j \neq k - N, \\ 0 & k = N + 1, \dots, N + M, j = k - N. \end{cases}$$

Then $Q = \cup_{k=1}^{N+M} Q_k$ with $Q_k = \{r \in \mathbb{R}^M \mid r < b_k\} = \{r \in \mathbb{R}^M \mid r_{k-N} < 0\}$ for $k = N + 1, \dots, N + M$. This will simplify theoretical considerations as we do not need to examine different cases.

Note that Q is not convex (see e.g. Figure 8.7) and hence we cannot use convex optimization algorithms to solve (8.10). However, Q has a very special structure that is the key in solving this problem. Therefore, let us now consider the connection between the solution of (8.10) and the function V . We prove the following two statements:

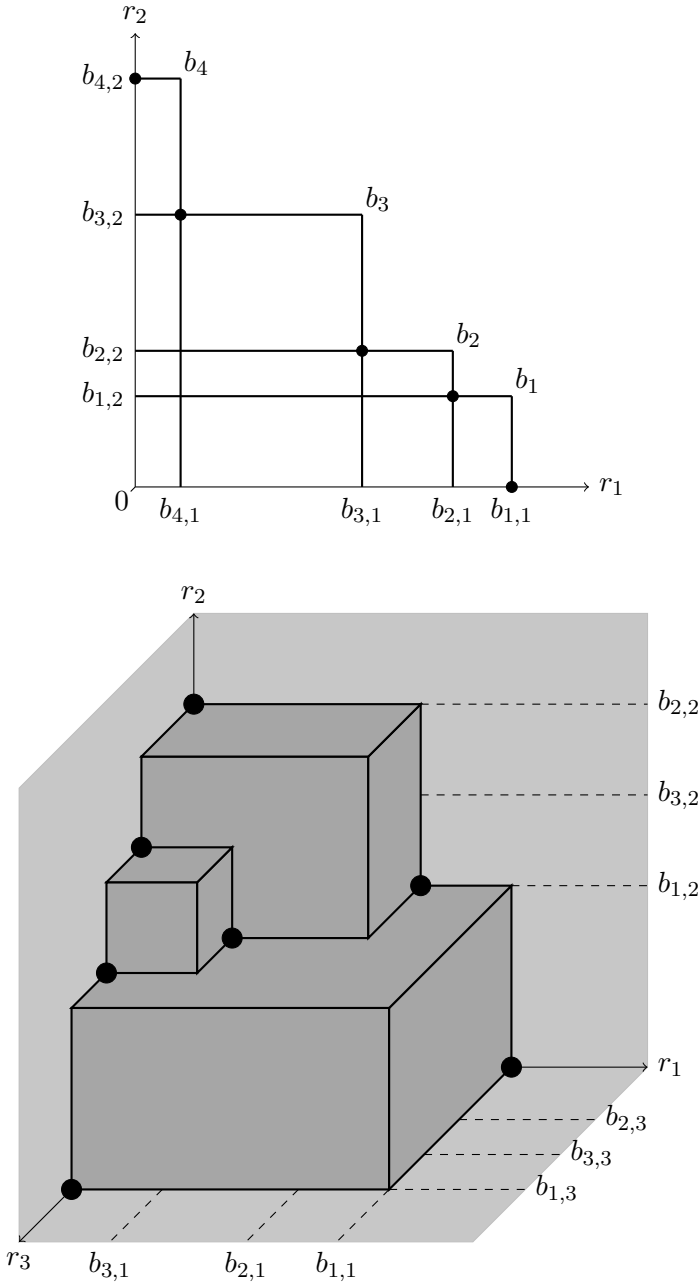


Fig. 8.7: The sets Q_k ($k = 1, \dots, N$) and the concave corners of Q (dots) with $M = 2, N = 4$ (top) and $M = N = 3$ (bottom).

Theorem 8.2.4:

Let $V : \mathbb{R}^M \rightarrow \mathbb{R}$ be (strictly) monotonically increasing on $\mathbb{R}^M \setminus Q$ in each component, i.e. $V(r) \leq V(r + \varepsilon e_j)$ for each $\varepsilon > 0$ and $j = 1, \dots, M$ where e_j is the j -th canonic unit vector. Then the concave corners of Q are the (only and strict) local minima of V on $\mathbb{R}^M \setminus Q$.

Proof:

Let r be a concave corner of Q and $\zeta \in \mathcal{S}$. Denote $\zeta_- = \min\{0, \zeta\}$ and $\zeta_+ = \max\{0, \zeta\}$ where min and max are evaluated component-wise. It follows that $\zeta = \zeta_- + \zeta_+$ and for $\varepsilon > 0$ and $\|\zeta_-\|_2 > 0$

$$r + \varepsilon \zeta = r + \underbrace{\varepsilon \zeta_-}_{\in Q} + \varepsilon \zeta_+.$$

Because Q is open we obtain that $r + \varepsilon \zeta \in Q$ for ε small enough. Hence for feasible directions $\zeta \in \mathcal{S}$ it holds that $\|\zeta_-\|_2 = 0$, i.e. $\zeta = \sum_{j=1}^M z_j e_j$ with $z_j \geq 0$ for all $j = 1, \dots, M$. Thus we obtain

$$V(r + \varepsilon \zeta) = V\left(r + \sum_{j=1}^M z_j e_j\right) \geq V\left(r + \sum_{j=2}^M z_j e_j\right) \geq \dots \geq V(r)$$

and consequential r is a local minimum.

Now consider V being strictly monotonic. We obtain that r is a strict local minimum in the same way. For $\tilde{r} \notin Q$ being no concave corner the existence of $\varepsilon > 0$ and e_j such that $r - \varepsilon e_j \notin Q$ follows by definition. Hence \tilde{r} is no local minimum because of $V(r) > V(r - \varepsilon e_j)$. \square

Corollary 8.2.5:

The following functions are monotonically increasing in each component on \mathbb{R}_+^M :

- All p -norms $\|r\|_p$ with $p \in [0, \infty]$.
- The two-dimensional volume function of the polygonal line

$$V(r) = \frac{1}{2} \sum_{P_j P_k \in \Delta(P_1, \dots, P_M)} r_j r_k \sqrt{1 - \langle \zeta_j, \zeta_k \rangle^2}.$$

- The three-dimensional volume function of the triangulation

$$V(r) = \frac{1}{3} \sum_{\Delta P_j P_k P_l \in \Delta(P_1, \dots, P_M)} r_j r_k r_l |\langle \zeta_3, \zeta_1 \times \zeta_2 \rangle|.$$

where $\zeta_1 \times \zeta_2$ denotes the cross product in \mathbb{R}^n .

Proof:

For $p \in (0, \infty)$ this follows from $r_j^p \leq (r_j + \varepsilon)^p$ for all $\varepsilon > 0$ and $r_j \geq 0$. For $p = 0$ this is easy to see because of $\|r + \varepsilon e_j\|_0 = \|r\|_0 + 1$ for $r_j = 0$ or $\|r + \varepsilon e_j\|_0 = \|r\|_0$ otherwise, and for $p = \infty$ we obtain $\|r + \varepsilon e_j\|_\infty = \max\{\max_{k \neq j} r_k, r_j + \varepsilon\} \geq \max\{\max_{k \neq j} r_k, r_j\} = \|r\|_\infty$.

The volume functions are linear with positive coefficients in each r_j and hence monotonically increasing. \square

As we see, the concave corners of Q are highly related to the solution of our minimization problem (8.10). Thereby, it seems to be less important which function V we want to minimize since the “most usual” functions all have their local minima at the same points. We assume from now on that the function V is monotonically increasing in each component and hence the minima are the concave corners of Q . We will return to the discussion of V at the end of this section to obtain some improvements of the basic algorithm. Much more important is the question, how the concave corners can be computed. Let us discuss the following statement:

Theorem 8.2.6:

Let the coefficients $\{b_{kj}\}_{k,j=1}^{N+M,M}$ from Definition 8.2.3 be given. Then it holds:

1. A concave corner r satisfies $\forall j : \exists k : r_j = b_{kj}$ and $r_l < b_{kl}$ for $l \neq j$.
2. For $r \in \mathbb{R}^M$ let $\forall j : \exists k : r_j = b_{kj}$ hold. Furthermore for $\{j \mid r_j > 0\}$ define the vectors $r^j \in \mathbb{R}^M$ with $r_i^j = r_i$ for $j \neq i$ and $r_j^j = \max_k \{b_{kj} \mid b_{kj} < r_j\}$. Note that the upper index corresponds to the vector while the lower index denotes the coefficient. Then r is a concave corner if and only if $r^j \in Q$ for $\{j \mid r_j > 0\}$.
3. For two concave corners $r \neq \tilde{r}$ neither $r \leq \tilde{r}$ nor $\tilde{r} \leq r$ is true.
4. Q has only a finite number of concave corners.

Proof:

The first statement is true since by definition $r - \varepsilon e_j \in Q$ for all $j = 1, \dots, M$ and hence $r - \varepsilon e_j < b_k$ for a $k \leq N + M$. It follows with $r \notin Q$ and thus $r \not< b_k$ that $r_j = b_{kj}$ and $r_l < b_{kl}$ for all $l \neq j$.

Let us consider the second statement. Note that since $r \in \mathbb{R}^M$ we have $r < \infty$ in each component. For a concave corner r it holds $r^j \in Q$ by definition since $r^j = r - \varepsilon e_j$ for an $\varepsilon > 0$. Now let $r^j \in Q$ for all $j \in \{j \mid r_j > 0\}$, i.e. there exists a k for each j such that $r^j < b_k$. We obtain for all $j \in \{j \mid r_j > 0\}$ that $b_{kj} > \max_{k'} \{b_{k'j} \mid b_{k'j} < r_j\}$ and hence $b_{kj} \geq r_j$. Since r and r^j only differ in the j -th component it follows that $r - \varepsilon e_j \in Q$ for all $\varepsilon > 0$. Moreover, for $j \in \{j \mid r_j = 0\}$ the vector $r - \varepsilon e_j$ is not in \mathbb{R}_+^M . Thus r is a concave corner.

For the next statement assume the contradiction. Given concave corners $r \neq \tilde{r}$ it should hold that $r \leq \tilde{r}$. Now it follows that there exists a $j \leq M$ with $r_j < \tilde{r}_j$ and hence $r \leq \tilde{r} - \varepsilon e_j \in Q$.

With the first statement we obtain that the concave corners are combinations of the finite coefficients of $\{b_{kj}\}_{k,j=1}^{N+M,M}$, i.e. $\{b_{kj}\}_{k,j=1}^{N,M} \cup \{0\}$, and hence their number is limited by $(N+1)^M$. \square

Thus, theoretically, we now have everything at hand to solve the problem (8.10). Although this minimization problem is non-convex, we have shown that the local minima of V form a finite set of points. Hence, the global minimum can be found by calculating all concave corners of Q and evaluating the function V at these points. Theorem 8.2.6 shows that the concave corners are a subset of the M -dimensional grid $\mathfrak{G} := \{x \in \mathbb{R}^M \mid \forall j : \exists k : x_j = b_{kj}\}$. Unfortunately, this set contains $(N+1)^M$ points. Therefore, the naive approach to test each point on the concave corner properties will have exponential runtime in M . This becomes even more crucial as $M \gg N$ in the case of unsupervised minimization. However, according to Theorem 8.2.6 we only need to evaluate several comparisons of the type $b_{kj}\{<, =, >\}b_{k'j}$ to determine if $r \in \mathfrak{G}$ is a concave corner. Use e.g. the fact that

$$r \text{ concave corner} \Leftrightarrow \forall j \in \{j \mid r_j > 0\} : r^j \in Q \quad (8.11)$$

$$\Leftrightarrow \forall j \in \{j \mid r_j > 0\} : \exists k : r^j < b_k. \quad (8.12)$$

Note that all $\{b_{kj}\}_{k,j=1}^{N+M,M}$ are known and thus the comparisons can be evaluated beforehand. As we need the comparisons between b_{kj} and $b_{k'j}$ for all $j \leq M$ and all $k, k' \leq N$ this is equivalent to sorting the sets $\{b_{kj}\}_{k=1}^N$ for each $j \leq M$. Further, we do not need to evaluate the comparisons for $k > N$ or $k' > N$ as these are trivial. Hence we have a numerical effort of $O(MN \log N)$. Once all the comparisons are determined, we only need to prompt the required comparisons. Nevertheless, the naive approach will fail again. A naive point based method that considers each point of the set \mathfrak{G} on its own will have exponential runtime in the number of requests. On the contrary, a request based method that only asks once for the comparison result of each pair $(b_{kj}, b_{k'j})$ and passes the information forward to each affected point, has exponential memory requirements as it has to save information for each point in \mathfrak{G} .

The spiderweb algorithm that we will introduce in this work is a point based method that uses the information obtained by analysing one point to successively reduce the number of possible concave corners. Inspired by equation (8.11) the method will not directly check if r is a concave corner, but only analyse if a point r is in Q or not. Furthermore, the algorithm will only return one concave corner, i.e. a local minimum of V and does not calculate the set of all concave corners to reconstruct the global minimum of V . In this way we are able to reduce the costs to a linear runtime. Thereby the spiderweb algorithm is based on

Theorem 8.2.7:

Given $\{b_{kj}\}_{k,j=1}^{N+M,M}$ with $b_k \neq (\infty)_{j=1}^M$ for each $k \leq N + M$ it holds:

Existence: Q has a concave corner.

Starting Point: The point $r \in \mathfrak{G}$ with $r_j = \max_k \{b_{kj} \mid b_{kj} < \infty\}$ is not in Q .

Test: If $r \notin Q$, then s^j with $s_i^j = r_i$ for $i \neq j$ and $s_j^j = \min_k \{b_{kj} \mid b_{kj} > r_j\}$ is no concave corner.

Memory effect: If $r - \varepsilon e_j \in Q$, then $\tilde{r} - \varepsilon e_j \in Q$ for $\tilde{r} \leq r$.

Descending direction: For $\tilde{r} \notin Q$ there exists a concave corner $r \leq \tilde{r}$.

Proof:

Existence: This follows from the Starting Point and the Descending direction property.

Starting Point: For $r \in Q$ there has to be a b_k with $r < b_k$. By definition of r it follows that $b_k = (\infty)_{j=1}^M$ what is the contradiction.

Test: This follows from (2) of Theorem 8.2.6.

Memory effect: For $r - \varepsilon e_j \in Q$ there exists b_k such that $r - \varepsilon e_j < b_k$. We obtain $\tilde{r} - \varepsilon e_j \leq r - \varepsilon e_j < b_k$ and hence $\tilde{r} - \varepsilon e_j \in Q$.

Descending direction: Let $\tilde{r} \notin Q$ and without loss of generality we assume that \tilde{r} is no concave corner. Then there exists a $j \leq M$ such that $\tilde{r} - \varepsilon e_j \notin Q$ for ε small enough. Since Q is open we can define $\varepsilon_j = \max\{\varepsilon > 0 \mid \tilde{r} - \varepsilon e_j\}$. Update $\tilde{r}^1 = \tilde{r} - \varepsilon_j e_j \notin Q$. By definition of ε_j it follows that $\tilde{r}^1 - \varepsilon e_j \in Q$ for all $\varepsilon > 0$. If \tilde{r}^1 is a concave corner, then we are done. Otherwise repeat the procedure for \tilde{r}^1 , i.e. there exists a $j^1 \leq M$ with $\tilde{r}^1 - \varepsilon e_{j^1} \notin Q$. Because of the memory effect $j^1 \neq j$ holds and hence we obtain a concave corner latest after M iterations. \square

Now we introduce

Algorithm 8.2.8 (Spiderweb algorithm):

Given $\{b_{kj}\}_{k,j=1}^{N+M,M}$ with $b_k \neq (\infty)_{j=1}^M$ for each $k \leq N + M$ perform the procedure:

Initialize $r \in \mathfrak{G}$ with $r_j = \max_k \{b_{kj} \mid b_{kj} < \infty\}$ for $j = 1, \dots, M$.

Set $I = \{1, \dots, M\}$.

Iterate

- If $I = \emptyset$ return r .
- Choose $j' \in I$ (randomly).
- Define \tilde{r} by $\tilde{r}_j = r_j$ for $j \neq j'$ and $\tilde{r}_{j'} = \max_k \{b_{kj'} \mid b_{kj'} < r_{j'}\}$.
- If $\tilde{r} \in Q$, then set $I = I \setminus \{j'\}$ else set $r = \tilde{r}$.

The spiderweb algorithm iteratively decreases the components of r until a concave corner is reached. The iteration in r is illustrated in Figure 8.8. If we visualize the set $\{P_j\}_{j=1}^M$ with $P_j = C + r_j \zeta_j$ for each iteration we obtain the typical spiderweb structure shown in Figure 8.5. Note that the algorithm chooses the descending direction at random. One may discuss also other criteria depending on the application. However, in our numerical

experiments the best results were achieved by choosing the direction randomly where the probability for each direction is equal. Before we go on to the next subsection to discuss the use of this algorithm in supervised minimization, we prove

Theorem 8.2.9:

The spiderweb algorithm calculates a concave corner in at most $O(MN)$ iterations.

Proof:

We use Theorem 8.2.7 to prove this theorem. Due to the starting point property the algorithm starts with a point outside Q . In each iteration we evaluate \tilde{r} by $\tilde{r}_j = r_j$ for $j \neq j'$ and $\tilde{r}_{j'} = \max_k \{b_{kj'} \mid b_{kj'} < r_{j'}\}$. If $\tilde{r} \notin Q$, then r is no concave corner due to the test property. On the other hand, if $\tilde{r} \in Q$ for a chosen $j' \in I$ then r might be a concave corner. Since $\tilde{r} \leq r$ holds, the algorithm produces a decreasing sequence in each component and hence we can remove j' from I due to (2) of Theorem 8.2.6 and the memory effect. It follows that if $I = \emptyset$ then $r - \varepsilon e_j \in Q$ for each $j = 1, \dots, M$ and thus r is a concave corner.

The number of coefficients $b_{kj} < \infty$ for each j is limited by $N+1$ (N coefficients obtained by the data and $b_{N+j,j} = 0$). This means, starting at $r \in \mathfrak{G}$ with $r_j = \max_k \{b_{kj} \mid b_{kj} < \infty\}$ for $j = 1, \dots, M$, we can decrease each component upto N times before it is set to 0 and consequently $r = 0$ after $O(MN)$ iterations. Since $r \notin Q$, $r - \varepsilon e_j = 0 - \varepsilon e_j \notin \mathbb{R}_+^M$ and hence $r - \varepsilon e_j \in Q$ this is a concave corner. \square

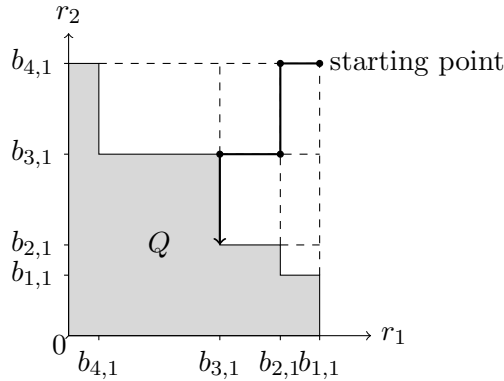


Fig. 8.8: The spiderweb algorithm starts at the “greatest” point of the grid and iterates downwards until a concave corner is reached, illustrated here for $N = 4$ and $M = 2$.

8.2.2 Supervised Minimization

We have shown how to find a local minimum of the function V on $\mathbb{R}^M \setminus Q$ with the help of the spiderweb algorithm. However, to deduce the discrete minimization problem (8.10) from the original reconstruction problem we used several assumptions. Thus, we now analyse the connection between the solutions of both problems. In this subsection we

will consider the supervised minimization problem where $\{\zeta_j\}_{j=1}^M$ or an approximation of it is known. Let us first analyse the case where $\{\zeta_j\}_{j=1}^M$ is exactly given. We can prove

Theorem 8.2.10:

Let $D \subset \mathbb{R}^n$ and $\{(t_k, h_k)\}_{k=1}^N$ given such that

$$ct_k = \sup_{P \in \partial D} (\|x_E(h_k) - P\|_2 + \|x_R(h_k) - P\|_2) \quad (8.13)$$

or

$$ct_k = \inf_{P \in \partial D} (\|x_E(h_k) - P\|_2 + \|x_R(h_k) - P\|_2) \quad (8.14)$$

holds, i.e. the measured signals result either from the nearest or most distant point. Note that these signals are the diffraction that is latest in time or the reflection that is earliest in time. Then it holds:

1. There exists a representation $\{P_j\}_{j=1}^M$ of ∂D that is also an irreducible solution for $\{(t_k, h_k)\}_{k=1}^N$.
2. Given $C \in \mathfrak{C}$, the directions $\{\zeta_j\}_{j=1}^M$ of this representation and Q, r with $r_j = \min_k \{b_{kj} \mid b_{kj} > 0\}$ is a concave corner of Q with $P_j = C + r_j \zeta_j$ for all $j = 1, \dots, M$.

Proof:

1. Since ∂D is closed and $\|\cdot\|_2$ is continuous, there exists a point $\tilde{P}_k \in \partial D$ with

$$ct_k = \|x_E(h_k) - \tilde{P}_k\|_2 + \|x_R(h_k) - \tilde{P}_k\|_2.$$

Hence $\{\tilde{P}_k\}_{k=1}^N$ is a representation of ∂D and a solution for $\{(t_k, h_k)\}_{k=1}^N$. If $\{\tilde{P}_k\}_{k=1}^N$ is not irreducible then there exists $k' \leq N$ such that $\{\tilde{P}_k\}_{k=1, k \neq k'}^N$ is also a solution. In this way we reduce the set $\{\tilde{P}_k\}_{k=1}^N$ until we obtain an irreducible set $\{P_j\}_{j=1}^M$.

2. Choose r with $r_j = \min_k \{b_{kj} \mid b_{kj} > 0\}$. Assume that $r \in Q$. It follows that $r < b_{k'}$ for some $k' \leq N$ and hence $b_{k'j} > \min_k \{b_{kj} \mid b_{kj} > 0\}$ for all $j \leq M$. Furthermore, following from Definition 8.2.2, there exists $j' \leq M$ and \tilde{k} such that $P_{j'} = C + s\zeta_{j'}$ with $s \geq b_{k'j'} > b_{\tilde{k}j'}$. Because $\overline{CP_{j'}}$ is in the interior of D (excluding $P_{j'}$), we obtain that $\tilde{P} = C + b_{\tilde{k}j'}\zeta_{j'}$ is in the interior of D . Moreover, this means that $E_{t_{\tilde{k}}, h_{\tilde{k}}} \cap (D \setminus \partial D)$ is non-empty and hence $E_{t, h_{\tilde{k}}} \cap (D \setminus \partial D)$ is non-empty for some $\varepsilon > 0$ and $|t - t_{\tilde{k}}| < \varepsilon$. Thus $(t_{\tilde{k}}, h_{\tilde{k}})$ does neither hold (8.13) nor hold (8.14).

Next we show that $P_j = C + r_j \zeta_j$. Therefore assume that $P_j = C + s\zeta_j$. By definition of b_{kj} it follows that $s \geq \min_k \{b_{kj} \mid b_{kj} > 0\} = r_j$. For $s > r_j$, i.e. $s > b_{kj}$

for some k , the point $P = C + b_{kj}\zeta_j$ is in the interior of D and hence as above (t_k, h_k) is not the first reflection or last diffraction. Thus $s = r_j$.

Now assume that r is not a concave corner. Following (2) in Theorem 8.2.6 there exists $j' \leq M$ with $r - r_{j'}e_{j'} \notin Q$ and hence $\{P_j\}_{j=1, j \neq j'}^M$ is also a solution for $\{t_k, h_k\}_{k=1}^N$. This is a contradiction to $\{P_j\}_{j=1}^M$ being irreducible.

□

As we see, for exactly given center point $C \in \mathfrak{C}$ and directions $\{\zeta_j\}_{j=1}^M$, we do not even need the spiderweb algorithm. The radii can be directly reconstructed using the formula $r_j = \min_k \{b_{kj} \mid b_{kj} > 0\}$. However, if $\{\zeta_j\}_{j=1}^M$ is only approximatively given, the reconstruction is not as simple as in the exact case. In the remaining part of this section we want to analyse how the concave corners of the approximated directions are related to the exact solution.

Before we introduce a more general definition for the coefficients b_{kj} , we anticipate a slightly confusion notation. In the following definition, we will introduce the sets $\widehat{\zeta}_j$ for $j \leq M$ and the set of sets $\{\widehat{\zeta}_j\}_{j=1}^M$. At first that might be irritating since ζ_j and the later introduced notation $\widetilde{\zeta}_j$ are directions, i.e. elements of the unit sphere $\mathcal{S} \subset \mathbb{R}^n$. However, we want to stay with this notation since ζ_j , $\widetilde{\zeta}_j$ and $\widehat{\zeta}_j$ are highly related as we will see soon. The reader should keep in mind that $\widehat{\zeta}_j$ denotes a set while ζ_j and $\widetilde{\zeta}_j$ are vectors, i.e. $\widehat{\zeta}_j \subset \mathcal{S}$ and $\zeta_j, \widetilde{\zeta}_j \in \mathcal{S}$. Figure 8.9 illustrates the sets $\{\widehat{\zeta}_j\}_{j=1}^5$, the original directions $\zeta_j \in \widehat{\zeta}_j$ for $j = 1, \dots, 5$ and the directions $\widetilde{\zeta}_j \in \widehat{\zeta}_j$ which will later be used as representatives for this set.

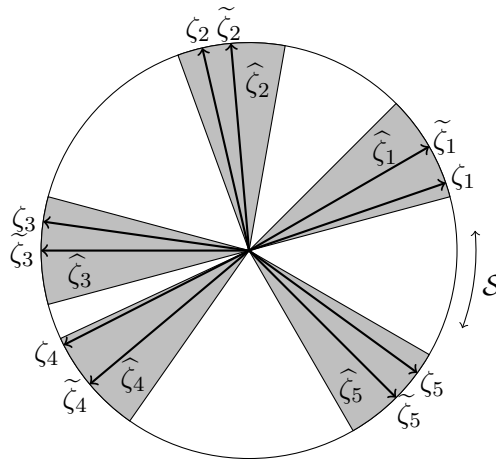


Fig. 8.9: The original (unknown) directions ζ_j , the sets $\widehat{\zeta}_j$ with $\zeta_j \in \widehat{\zeta}_j$ (gray) and arbitrary chosen representatives $\widetilde{\zeta}_j \in \widehat{\zeta}_j$ for $j = 1, \dots, 5$ in $\mathcal{S} \subset \mathbb{R}^2$.

Definition 8.2.11:

Let $\{\widehat{\zeta}_j\}_{j=1}^M$ be M subsets of $\mathcal{S} = \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$, i.e. $\widehat{\zeta}_j \subset \mathcal{S}$ for all $j = 1, \dots, M$. For $\varepsilon > 0$ we call $\{\widehat{\zeta}_j\}_{j=1}^M$ a *separation* if $\widehat{\zeta}_j \cap \widehat{\zeta}_{j'} = \emptyset$ for $j \neq j'$. Furthermore, we call $\{\widehat{\zeta}_j\}_{j=1}^M$ an ε -*separation* if it is a separation and

$$\max_{\zeta, \zeta' \in \widehat{\zeta}_j} \|\zeta - \zeta'\|_2 \leq \varepsilon.$$

We call $\{\widehat{\zeta}_j\}_{j=1}^M$ *complete* if

$$\bigcup_{j=1}^M \widehat{\zeta}_j = \mathcal{S}.$$

otherwise we say that $\{\widehat{\zeta}_j\}_{j=1}^M$ is *incomplete*. Now for given $C \in \mathfrak{C}$, $\{(t_k, h_k)\}_{k=1}^N$ and $k = 1, \dots, N$ define the sets $B_k(\zeta) := \{r \mid C + r\zeta \in E_{t_k, h_k}\}$ and

$$\widetilde{b}_{kj} = \inf_{\zeta \in \widehat{\zeta}_j} (\min \{(B_k(\zeta) \cap \mathbb{R}_+) \cup \{\infty\}\}).$$

Furthermore with $\widetilde{b}_k = (\widetilde{b}_{kj})_{j=1}^M$ define $\widetilde{Q}_k := \{r \in \mathbb{R}^M \mid r < \widetilde{b}_k\}$ and $\widetilde{Q} := \bigcup_{k=1}^N \widetilde{Q}_k \cup (\mathbb{R}^M \setminus \mathbb{R}_+^M)$.

We first note that this is a more general definition as Definition 8.2.2 and 8.2.3 since $\widetilde{b}_{kj} = b_{kj}$ and hence $\widetilde{Q} = Q$ for $\widehat{\zeta}_j = \{\zeta_j\}$ for $j \leq M$. For a better understanding of Definition 8.2.11 and the aim of this subsection, we introduce the following considerations.

Let the directions $\{\zeta_j\}_{j=1}^M$ be given approximatively by the directions $\{\widetilde{\zeta}_j\}_{j=1}^M$, i.e. $\|\zeta_j - \widetilde{\zeta}_j\|_2 \leq \varepsilon$ for an $\varepsilon > 0$. Then $\zeta_j \in B_\varepsilon(\widetilde{\zeta}_j)$ where $B_\varepsilon(\widetilde{\zeta}_j) = \{\zeta \in \mathcal{S} \mid \|\zeta - \widetilde{\zeta}_j\|_2 \leq \varepsilon\}$. Hence, we could also say that the sets $B_\varepsilon(\widetilde{\zeta}_j)$ for $j = 1, \dots, M$ with $\zeta_j \in B_\varepsilon(\widetilde{\zeta}_j)$ are given or more general: For $\{\zeta_j\}_{j=1}^M$ let the sets $\{\widehat{\zeta}_j\}_{j=1}^M$ with $\zeta_j \in \widehat{\zeta}_j$ be given. Assuming now that the sets $\{\widehat{\zeta}_j\}_{j=1}^M$ have special properties, we can show some approximation results as e.g. the connection between the concave corners of Q and \widetilde{Q} . As the sets $B_\varepsilon(\widetilde{\zeta}_j)$ will normally not have these properties, the analysis we will present for the supervised minimization is more of theoretical nature. However, for unsupervised minimization we will design a complete ε -separation that has those properties. Because this ε -separation covers \mathcal{S} , there exists a subset of sets $\{\widehat{\zeta}_j\}_{j=1}^M$ such that $\zeta_j \in \widehat{\zeta}_j$ holds and thus we can apply the analysis of supervised minimization. To analyse the concave corners of \widetilde{Q} we need the following definition.

Definition 8.2.12:

Let $\{\widehat{\zeta}_j\}_{j=1}^M$ be a (complete or incomplete) separation. We say that $\{\widehat{\zeta}_j\}_{j=1}^M$ is *adapted* to $\{t_k, h_k\}_{k=1}^N$ if for all $k \leq N$ and all $j \leq M$ one of the conditions

- a) $B_k(\zeta) \cap \mathbb{R}_+^M = \emptyset$ for all $\zeta \in \widehat{\zeta}_j$
- b) $B_k(\zeta) \cap \mathbb{R}_+^M \neq \emptyset$ for all $\zeta \in \widehat{\zeta}_j$

holds. This means there is either for no direction or for all directions $\zeta \in \widehat{\zeta}_j$ an intersection between the ray $\{C + r\zeta \mid r \geq 0\}$ and the spheroid E_{t_k, h_k} . To avoid notational clutter, we also say that $\{\widehat{\zeta}_j\}_{j=1}^M$ is *adapted to the measurements* if we consider the case of arbitrary given $\{t_k, h_k\}_{k=1}^N$.

The sets defined in both definitions are illustrated in Figure 8.10. Here an incomplete ε -separation as well as a complete separation that is adapted to the measurements is shown for the two-dimensional case. With the help of these definitions we are able to prove our first result.

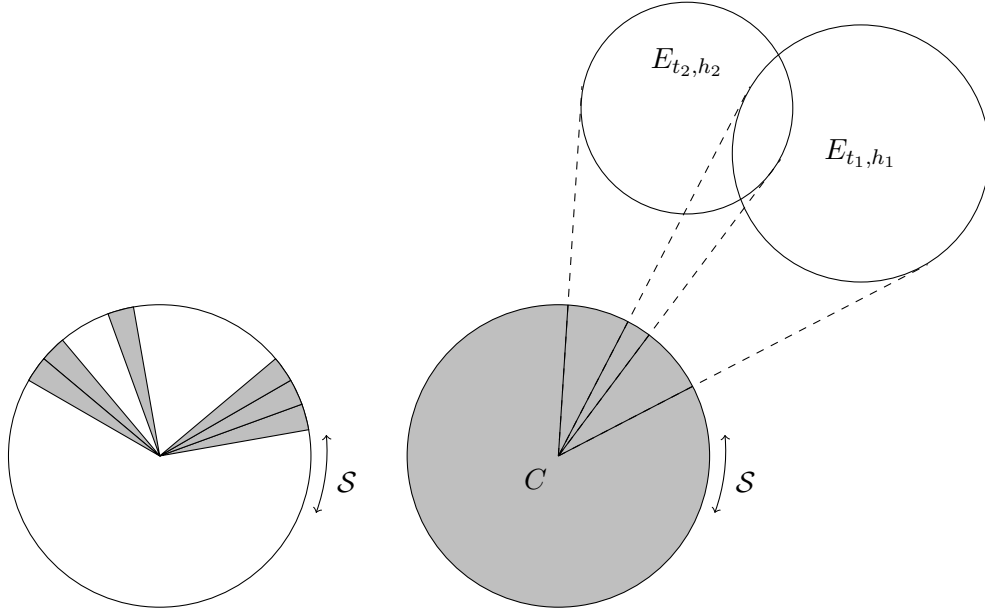


Fig. 8.10: An incomplete ε -separation with 6 sets that does not fill the complete sphere \mathcal{S} (left); an complete separation with 4 sets that is adapted to $\{(t_1, h_1), (t_2, h_2)\}$, here the center of \mathcal{S} is set to the center point C to illustrate the connection to $E_{t_1, h_1}, E_{t_2, h_2}$ (right).

Theorem 8.2.13:

For a center point C let $\{\zeta_j\}_{j=1}^M$ be the directions of the exact solution and for $\varepsilon > 0$ let $\{\widehat{\zeta}_j\}_{j=1}^M$ be a (complete or incomplete) ε -separation adapted to $\{\mathbf{t}_k, h_k\}_{k=1}^N$ with $\zeta_j \in \widehat{\zeta}_j$ for all $j = 1, \dots, M$. Then there exists $\widetilde{\varepsilon} > 0$ such that

$$b_{kj} - \widetilde{b}_{kj} \leq \widetilde{\varepsilon}$$

holds for all $k = 1, \dots, N$ and $j = 1, \dots, M$ with $\widetilde{b}_{kj} < \infty$. Furthermore, it holds that

$$\lim_{\varepsilon \rightarrow 0} \widetilde{\varepsilon} = 0.$$

Proof:

Since $\{\widehat{\zeta}_j\}_{j=1}^M$ is adapted to $\{\mathbf{t}_k, h_k\}_{k=1}^N$ it follows that $b_{kj} < \infty$ for $\widetilde{b}_{kj} < \infty$. We define the sets

$$Y_k := \{\zeta \in \mathcal{S} \mid \exists r > 0 : C + r\zeta \in E_{\mathbf{t}_k, h_k}\}$$

and the function

$$y_k : Y_k \rightarrow \mathbb{R}_+ \quad , \quad y_k(\zeta) = \min\{r > 0 \mid C + r\zeta \in E_{\mathbf{t}_k, h_k}\}.$$

It directly follows that for each $j = 1, \dots, M$ there exists a $k \leq N$ with $\widehat{\zeta}_j \subset Y_k$. Moreover $b_{kj} = y_k(\zeta_j)$ and $\widetilde{b}_{kj} = y_k(\widetilde{\zeta}_j)$ with some $\widetilde{\zeta}_j \in \widetilde{\widehat{\zeta}_j}$ where $\widetilde{\widehat{\zeta}_j}$ is the closure of $\widehat{\zeta}_j$. Since $E_{\mathbf{t}_k, h_k}$ is continuous, we obtain that y_k is continuous and furthermore it is uniformly continuous because Y_k is compact. Since $\widetilde{\zeta}_j, \zeta_j \in \widetilde{\widehat{\zeta}_j}$ we obtain $\|\widetilde{\zeta}_j - \zeta_j\|_2 \leq \varepsilon$ and with the uniform continuity it follows the existence of $\widetilde{\varepsilon} > 0$ such that

$$b_{kj} - \widetilde{b}_{kj} < \widetilde{\varepsilon}$$

holds where $\lim_{\varepsilon \rightarrow 0} \widetilde{\varepsilon} = 0$ also follows from the uniform continuity. \square

Note that in the general case the connection between ε and $\widetilde{\varepsilon}$ is quite complicated and cannot be given by an exact analytical expression. However, during the reconstruction process the function y_k has to be analysed anyway since $\widetilde{b}_{kj} = \inf_{\zeta \in \widehat{\zeta}_j} y_k(\zeta)$. From

$$b_{kj} - \widetilde{b}_{kj} \leq \sup_{\zeta \in \widehat{\zeta}_j} y_k(\zeta) - \widetilde{b}_{kj}$$

we can derive numerically the upper bound

$$\widetilde{\varepsilon} \leq \max_{k \leq N} \max_{\substack{j \leq M \\ b_{kj} \neq \infty}} \left(\sup_{\zeta \in \widehat{\zeta}_j} y_k(\zeta) - \inf_{\zeta \in \widehat{\zeta}_j} y_k(\zeta) \right) \quad (8.15)$$

during the process. Moreover, this estimate is sharp if the exact directions $\{\zeta_j\}_{j=1}^M$ are unknown since for indices k, j where the maximum of (8.15) is reached, y_k is continuous and hence there exists a sequence $\{z_l\}_{l=1}^\infty \subset \widehat{\zeta}_j$ such that

$$\lim_{l \rightarrow \infty} y_k(z_l) = \sup_{\zeta \in \widehat{\zeta}_j} y_k(\zeta).$$

For $\varepsilon' < \widetilde{\varepsilon}$, it follows that

$$\lim_{l \rightarrow \infty} y_k(z_l) - \widetilde{b}_{kj} = \widetilde{\varepsilon} > \varepsilon'.$$

Now choose $\zeta_j = z_l$ with l large enough, then it follows that

$$b_{kj} - \widetilde{b}_{kj} = y_k(z_l) - \widetilde{b}_{kj} > \varepsilon'.$$

Hence Theorem 8.2.13 holds for

$$\widetilde{\varepsilon} = \max_{k \leq N} \max_{\substack{j \leq M \\ b_{kj} \neq \infty}} \left(\sup_{\zeta \in \widehat{\zeta}_j} y_k(\zeta) - \inf_{\zeta \in \widehat{\zeta}_j} y_k(\zeta) \right).$$

Note that this is a worst case estimation and $\widetilde{\varepsilon}$ might be much smaller depending on the unknown exact directions $\{\zeta_j\}_{j=1}^M$. Hence the approximation error will often be smaller in applications as $\widetilde{\varepsilon}$ may convey.

With the results obtained above we are able to analyse the connection between the concave corners of \widetilde{Q} and the exact solution (2) in Theorem 8.2.10. This will be done in the following two theorems. Given a concave corner $\widetilde{r} = (\widetilde{r}_j)_{j=1}^M$ the first theorem will provide a sufficient condition for \widetilde{r} being an approximation of the exact solution r . In the second theorem we will prove the existence of a concave corner that satisfies the sufficient condition.

Theorem 8.2.14:

For a given center point C let $r = (r_j)_{j=1}^M$ and $\{\zeta_j\}_{j=1}^M$ be the radii and directions of the exact solution. Moreover, for $\varepsilon > 0$ let $\{\widehat{\zeta}_j\}_{j=1}^M$ be a (complete or incomplete) ε -separation with $\zeta_j \in \widehat{\zeta}_j$ that is adapted to the measurements. Now let $\widetilde{r} = (\widetilde{r}_j)_{j=1}^M$ be a concave corner of \widetilde{Q} and $\widetilde{\varepsilon} > 0$ be given by Theorem 8.2.13. If

$$\widetilde{r}_j - \widetilde{b}_{kj} \leq \widetilde{\varepsilon}$$

holds for all $k = 1, \dots, N$ and $j = 1, \dots, M$ (where we define $\widetilde{r}_j - \infty = -\infty$), then

$$\|\widetilde{r} - r\|_\infty \leq \widetilde{\varepsilon}$$

and hence \widetilde{r} is an approximation of r in the maximum-norm.

Proof:

We will prove this theorem for the j -th component where $j = 1, \dots, M$ is arbitrarily chosen. The result then follows directly for the complete vector. With (2) in Theorem 8.2.10 we obtain

$$r_j + \tilde{\varepsilon} = \min_{k \leq N} b_{kj} + \tilde{\varepsilon} \geq \min_{k \leq N} \tilde{b}_{kj} + \tilde{\varepsilon} \geq \tilde{r}_j.$$

Furthermore, there exists a k such that with Theorem 8.2.13 it follows that

$$\tilde{r}_j = \tilde{b}_{kj} \geq b_{kj} - \tilde{\varepsilon} \geq r_j - \tilde{\varepsilon}$$

because $r_j = \min_{k' \leq N} b_{k'j} \leq b_{kj}$. □

Theorem 8.2.15:

Using the notations of Theorem 8.2.14, there always exists a concave corner \tilde{r} of \tilde{Q} such that

$$\tilde{r}_j - \tilde{b}_{kj} \leq \tilde{\varepsilon}$$

holds for all $k = 1, \dots, N$ and $j = 1, \dots, M$.

Proof:

Let r be the exact solution. Since $r \notin Q$, for each $j = 1, \dots, M$ there exists a $k \leq N$ such that

$$r_j \geq b_{kj} \geq \tilde{b}_{kj}$$

and hence $r \notin \tilde{Q}$. Now consider the point $x = (x_j)_{j=1}^M$ with $x_j = \min_{k \leq N} \tilde{b}_{kj} + \tilde{\varepsilon}$. It follows that

$$x_j = \min_{k \leq N} \tilde{b}_{kj} + \tilde{\varepsilon} \geq \min_{k \leq N} b_{kj} = r_j$$

and thus $x \notin \tilde{Q}$. Following from the memory effect of Theorem 8.2.7 there exist a concave corner \tilde{r} of \tilde{Q} with $\tilde{r} \leq x$. We obtain for all $j = 1, \dots, M$

$$\tilde{r}_j - \tilde{b}_{kj} \leq x_j - \tilde{b}_{kj} = \min_{k \leq N} \tilde{b}_{kj} + \tilde{\varepsilon} - \tilde{b}_{kj} \leq \tilde{\varepsilon}.$$

□

With these results we are now able to design an algorithm that solves the supervised minimization problem. Because this algorithm will need a specially designed (complete or incomplete) ε -separation, it is most unlikely that one can use it in practice. However, the algorithm is the basis for our reconstruction method for the unsupervised minimization and hence we will consider it here.

Algorithm 8.2.16 (Supervised minimization):

Given the measurements $\{t_k, h_k\}_{k=1}^N$ and a (complete or incomplete) ε -separation that is adapted to the measurements, perform the following procedure

1. Calculate $\tilde{b}_{kj} = \inf_{\zeta \in \hat{\zeta}_j} (\min \{(B_k(\zeta) \cap \mathbb{R}_+) \cup \{\infty\}\})$.

2. Initialize $\tilde{\varepsilon} := \max_{k \leq N} \max_{\substack{j \leq M \\ b_{kj} \neq \infty}} \left(\sup_{\zeta \in \hat{\zeta}_j} y_k(\zeta) - \inf_{\zeta \in \hat{\zeta}_j} y_k(\zeta) \right)$.

3. Use the spiderweb algorithm 8.2.8 to reconstruct a concave corner \tilde{r} of \tilde{Q} where the initialization step is changed to:

Initialize $r \in \mathfrak{G}$ with $r_j = \max_k \{\tilde{b}_{kj} \mid \tilde{b}_{kj} \leq \min_k \tilde{b}_{kj} + \tilde{\varepsilon}\}$ for $j = 1, \dots, M$.

Here, the change of the initialization of the spiderweb algorithm will not only ensure that the determined concave corner holds the approximation property $\|\tilde{r} - r\|_\infty \leq \tilde{\varepsilon}$ where r is the exact solution, but it also reduces the number of performed iterations since the starting point is nearer to the concave corner as the point chosen before. In Figure 8.5 we illustrated the iterations of the spiderweb algorithm where we used the starting point

$$r_j = \max_k \{\tilde{b}_{kj} \mid \tilde{b}_{kj} < \infty\}$$

on the left side and

$$r_j = \max_k \{\tilde{b}_{kj} \mid \tilde{b}_{kj} \leq \min_k \tilde{b}_{kj} + \tilde{\varepsilon}\}$$

on the right side for the same defect. As one can see, the left figure is much more zoomed out due to the fact that the iteration starts at a point that is far away from the solution. The different sizes of both reconstructions are caused by the different scales of both figure parts. We can prove

Theorem 8.2.17:

Let r' be the solution for the exact directions $\{\zeta_j\}_{j=1}^M$. Then Algorithm 8.2.16 returns a concave corner \tilde{r} with

$$\|r' - \tilde{r}\|_\infty \leq \tilde{\varepsilon}.$$

Proof:

For the initialization $r \in \mathfrak{G}$ with $r_j = \max_k \{\tilde{b}_{kj} \mid \tilde{b}_{kj} \leq \min_k \tilde{b}_{kj} + \tilde{\varepsilon}\}$ for $j = 1, \dots, M$ the existence of a concave corner $\tilde{r} \in \mathfrak{G}$ with $\tilde{r} \leq r$ follows by Theorem 8.2.10 and hence $r \notin Q$. Thus, the initialization for the spiderweb algorithm is correct and the algorithm will return a concave corner \tilde{r} . Since

$$\tilde{r}_j - \tilde{b}_{kj} \leq r_j - \tilde{b}_{kj} \leq r_j - \min_k \tilde{b}_{kj} \leq \tilde{\varepsilon}$$

the statement follows by Theorem 8.2.14. □

Hence, we now have an algorithm to solve the supervised minimization problem. The algorithm may not be applicable in practise as we need an ε -separation that is adapted to the measurements. However, we will extend the procedure to the unsupervised minimization problem. Here it will be possible to chose the ε -separation and thus it can be designed in a way such that all conditions hold. Before we continue with the next subsection, let us point out another result that estimates the total error between the original representation $\{P_j\}_{j=1}^M$ and its approximation $\{\tilde{P}_j\}_{j=1}^M$.

Theorem 8.2.18:

For measurements $\{t_k, h_k\}_{k=1}^N$ and a center point $C \in \mathfrak{C}$ let r and $\{\zeta_j\}_{j=1}^M$ be the exact radii and directions of the representation $\{P_j\}_{j=1}^M$, i.e. $P_j = C + r_j \zeta_j$ for all $j = 1, \dots, M$. Furthermore, for an ε -separation $\{\hat{\zeta}_j\}_{j=1}^M$ with $\zeta_j \in \hat{\zeta}_j$ that is adapted to the measurements let \tilde{r} be the approximated radii. Now choose $\tilde{\zeta}_j \in \hat{\zeta}_j$ then it follows that

$$\|P_j - \tilde{P}_j\|_2 \leq \tilde{\varepsilon} + \tilde{r}_j \varepsilon$$

with $\tilde{P}_j = C + \tilde{r}_j \tilde{\zeta}_j$.

Proof:

$$\begin{aligned} \|P_j - \tilde{P}_j\|_2 &= \|C + r_j \zeta_j - C - \tilde{r}_j \tilde{\zeta}_j\|_2 \\ &= \|r_j \zeta_j - \tilde{r}_j \zeta_j + \tilde{r}_j \zeta_j - \tilde{r}_j \tilde{\zeta}_j\|_2 \\ &\leq |r_j - \tilde{r}_j| + \tilde{r}_j \|\zeta_j - \tilde{\zeta}_j\|_2 \\ &\leq \tilde{\varepsilon} + \tilde{r}_j \varepsilon \end{aligned}$$

□

8.2.3 Unsupervised Minimization

In the last subsection we have discussed the case of supervised minimization in detail. Now we want to consider the unsupervised case where no approximation of the directions $\{\zeta_j\}_{j=1}^M$ is given. Therefore, we will analyse two considerations. First, we will discuss the connection between the concave corners of two sets Q^1, Q^2 where both sets were obtained considering the same measurements but Q^1 is obtained using the incomplete ε -separation $\{\hat{\zeta}_j\}_{j=1}^M$ while Q^2 uses an ε -separation with more sets $\{\hat{\zeta}_j\}_{j=1}^{M'}$ where $M' > M$. We will see that the concave corners of Q^1 directly correlate with concave corners of Q^2 . This result will allow us to use more directions as needed in the beginning and select a subset of them at the end. Although we do not know the exact directions, we know that they are contained in a subset of the ε -separation. Hence, we can start with a complete ε -separation, choose a subset i.e. an incomplete ε -separation and apply the theory obtained for the supervised minimization. The second result of this subsection will give

a guidance, how one can design an ε -separation that is adapted to the measurements. Therefore, we will compose it out of two sets where one is an ε -separation and the other is adapted to the measurements. Combining both results we will be able to adapt Algorithm 8.2.16 to the unsupervised case. Let us now consider the first result, namely

Theorem 8.2.19:

Let the measurements $\{t_k, h_k\}$ and the center point $C \in \mathfrak{C}$ be given. Furthermore, let Q^1 be obtained considering the incomplete ε -separation $\{\hat{\zeta}_j\}_{j=1}^M$ and Q^2 is obtained using an ε -separation with more sets $\{\hat{\zeta}_j\}_{j=1}^{M'}$ with $M' > M$ where the first M sets for Q^1 and Q^2 are equal. Then $r^1 \in \mathbb{R}^M$ is a concave corner of Q^1 if and only if

$$r^2 = \begin{pmatrix} r^1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{M'}$$

is a concave corner of Q^2 .

Proof:

For $k = 1, \dots, N$ we denote by $\tilde{b}_k^1 \in \mathbb{R}^M$ and $\tilde{b}_k^2 \in \mathbb{R}^{M'}$ the vectors defined in Definition 8.2.11 belonging to Q^1 respectively Q^2 . By definition it is clear that

$$\tilde{b}_k^2 = \begin{pmatrix} \tilde{b}_k^1 \\ x_k \end{pmatrix} \in \mathbb{R}^{M'}$$

with $x_k \in \mathbb{R}^{M-M'}$ and $x_k > 0$. Now let $r^1 \in \mathbb{R}^M$ be a concave corner of Q^1 , this especially means $r^1 \not\prec \tilde{b}_k^1$ holds for all $k = 1, \dots, N$. Then for

$$r^2 = \begin{pmatrix} r^1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{M'}$$

it follows that $r^2 \not\prec \tilde{b}_k^2$ for all $k = 1, \dots, N$. Thus $r^2 \notin Q^2$. Furthermore for $\varepsilon' > 0$ it holds that $r^2 - \varepsilon' e_l \in Q^2$ for $l > M$ because it will contain a negative component where e_l is the l -th unit vector of dimension M' . Since r^1 is a concave corner, for each $l \leq M$ there exists a $k \leq N$ such that $r^1 - \varepsilon' e_l < \tilde{b}_k^1$ where e_l is the l -th unit vector of dimension M . It follows that also $r^2 - \varepsilon' e_l < \tilde{b}_k^2$ and moreover r^2 is a concave corner of Q^2 .

Now let

$$r^2 = \begin{pmatrix} r^1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{M'}$$

be a concave corner of Q^2 . This means for all $l \leq M$ and $\varepsilon' > 0$ there exists a $k \leq N$ such that $r^2 - \varepsilon' e_l < \tilde{b}_k^2$. Considering only the first M inequalities we obtain $r^1 - \varepsilon' e_l < \tilde{b}_k^1$ and hence r^1 is a concave corner of Q^1 . \square

Note that Theorem 8.2.19 also holds if the additional sets are not added at the end, e.g. Q^1 uses the ε -separation $\{\zeta_{2j}\}_{j=1}^M$ while Q^2 uses $\{\zeta_j\}_{j=1}^{2M}$. We just used this presentation form for simplicity. For our next result we need the following definition.

Definition 8.2.20:

Let $A = \{A_i\}_{i=1}^{M_1}$ and $B = \{B_j\}_{j=1}^{M_2}$ be two complete separations of \mathcal{S} . We call

$$A \sqcap B := \left\{ Y \subset \mathcal{S} \left| \begin{array}{l} \exists i \leq M_1, j \leq M_2 : Y \subset A_i \cap B_j, \\ Y \text{ connected,} \\ \nexists \zeta \in \mathcal{S} \setminus Y : (Y \cup \{\zeta\}) \subset (A_i \cap B_j) \text{ is connected} \end{array} \right. \right\}.$$

the *common refinement* of A and B , i.e. $A \sqcap B$ is the set of all connected components of intersections between sets of A and B .

Figure 8.11 shows an example for the common refinement of an ε -separation and a complete separation that is adapted to some measurements (see also Figure 8.10). We will use such refinements in our algorithm. The following theorem will prove that both properties, ε -separation and being adapted to the measurements, will be preserved when building a common refinement.

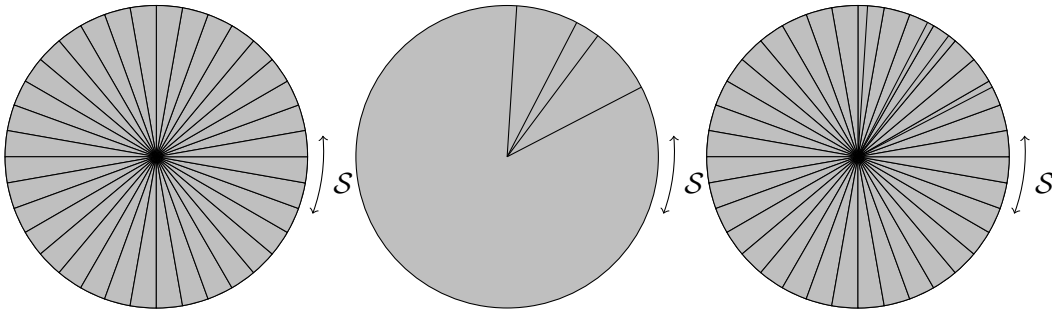


Fig. 8.11: A complete ε -separation (left), a separation that is adapted to some measurements (middle), see also Figure 8.10, and their common refinement (right).

Theorem 8.2.21:

Let $A = \{A_i\}_{i=1}^{M_1}$ and $B = \{B_j\}_{j=1}^{M_2}$ be two complete separations of \mathcal{S} . Then it holds:

- $A \sqcap B$ is a complete separation of \mathcal{S} .
- If A or B is an ε -separation, then $A \sqcap B$ is an ε -separation.
- If A or B is adapted to $\{\mathbf{t}_k, h_k\}_{k=1}^N$, then $A \sqcap B$ is adapted to the measurements.

Proof:

- For each $\zeta \in \mathcal{S}$ there exists $i \leq M_1$ and $j \leq M_2$ such that $\zeta \in A_i$ and $\zeta \in B_j$. It follows that $\zeta \in A_i \cap B_j$, hence there exists $Y \in A \sqcap B$ such that $\zeta \in Y$ and hence

$$\bigcup_{Y \in A \sqcap B} Y = \mathcal{S}.$$

For $Y_1, Y_2 \in A \sqcap B$ with $Y_1 \neq Y_2$ there exists $i, i' \leq M_1$ and $j, j' \leq M_2$ such that

$$\begin{aligned} Y_1 &\subset A_i \cap B_j \\ Y_2 &\subset A_{i'} \cap B_{j'}. \end{aligned}$$

If $i \neq i'$ or $j \neq j'$ it follows that

$$Y_1 \cap Y_2 = (A_i \cap B_j) \cap (A_{i'} \cap B_{j'}) = (A_i \cap A_{i'}) \cap (B_j \cap B_{j'}) = \emptyset.$$

For $i = i'$ and $j = j'$ assume $Y_1 \cap Y_2 \neq \emptyset$. Then $Y_1 \cup Y_2 \subset A_i \cap B_j$ is connected. Thus $Y_1, Y_2 \notin A \sqcap B$ by the third condition of the definition. Hence $A \sqcap B$ is a complete separation of \mathcal{S} .

- Without loss of generality let A be an ε -separation. Now choose $Y \in A \sqcap B$ with $Y \subset A_i \cap B_j$. Then for all $\zeta, \zeta' \in Y$ it holds that $\|\zeta - \zeta'\|_2 \leq \varepsilon$ because $\zeta, \zeta' \in A_i$. Thus $A \sqcap B$ is an ε -separation.
- Without loss of generality let B be adapted to $\{\mathbf{t}_k, h_k\}_{k=1}^N$. This means for all $k \leq N$ and all $j \leq M_2$ one of the conditions

- a) $B_k(\zeta) \cap \mathbb{R}_+^M = \emptyset$ for all $\zeta \in B_j$
- b) $B_k(\zeta) \cap \mathbb{R}_+^M \neq \emptyset$ for all $\zeta \in B_j$

holds. Then $A \sqcap B$ is also adapted to the measurements since the conditions also hold for $\zeta \in A_i \cap B_j$.

□

With the above obtained results, we are able to design an algorithm that solves the unsupervised minimization problem. Because this will be the most common case in practice, we directly present an expanded version of the algorithm that not only reconstructs a concave corner but also includes the reconstruction of $C \in \mathfrak{C}$ and the points $\{\tilde{P}_j\}_{j=1}^M$.

Algorithm 8.2.22 (Defect reconstruction):

Given the measurements $\{t_k, h_k\}_{k=1}^N$ perform the following reconstruction steps:

1. Calculate $C \in \mathfrak{C}$ with e.g. one of the heuristic methods presented in Section 8.1.
2. Choose $\varepsilon > 0$ and determine a complete ε -separation A of \mathcal{S} .
3. Determine a set of sets B that is adapted to the measurements $\{t_k, h_k\}_{k=1}^N$.
4. Set $\{\widehat{\zeta}_j\}_{j=1}^M = A \cap B$ where $M \gg N$ should hold.
5. Choose arbitrary points $\{\widetilde{\zeta}_j\}_{j=1}^M$ with $\widetilde{\zeta}_j \in \widehat{\zeta}_j$ for all $j \leq M$.
6. Calculate $\widetilde{b}_{kj} = \inf_{\zeta \in \widehat{\zeta}_j} (\min \{(B_k(\zeta) \cap \mathbb{R}_+) \cup \{\infty\}\})$.
7. Initialize $\widetilde{\varepsilon} := \max_{k \leq N} \max_{j, \widetilde{b}_{kj} \neq \infty} \left(\sup_{\zeta \in \widehat{\zeta}_j} y_k(\zeta) - \inf_{\zeta \in \widehat{\zeta}_j} y_k(\zeta) \right)$.
8. Use the spiderweb algorithm 8.2.8 to reconstruct a concave corner \widetilde{r} of \widetilde{Q} where the initialization step is changed to:
Initialize $r \in \mathfrak{G}$ with $r_j = \max_k \{\widetilde{b}_{kj} \mid \widetilde{b}_{kj} \leq \min_k \widetilde{b}_{kj} + \widetilde{\varepsilon}\}$ for $j = 1, \dots, M$.
9. Remove all 0-radii, i.e. set $\widetilde{r} = (\widetilde{r}_j)_{j=1, \widetilde{r}_j \neq 0}^M \in \mathbb{R}^{M'}$ with $M' \leq M$.
10. Return $\{\widetilde{P}_j\}_{j=1}^{M'}$ with $\widetilde{P}_j = C + \widetilde{r}_j \widetilde{\zeta}_j$.

As we will see next in this section, the algorithm provides an approximate solution for the defect reconstruction problem. However, the proof will be different to the previous results. While in earlier results we have assumed that the directions were approximately given and thus we are able to give approximation properties according to the exact radii, this is no longer possible for the unsupervised case. The problem here is not only that the exact directions are unknown. Moreover, since the reconstruction problem has no unique solution, there may exist more than one exact solution e.g. $\{P_j\}_{j=1}^M$ with $P_j = C + r_j \zeta_j$ and $\{P_l\}_{l=1}^{M'}$ with $P_l = C + r'_l \zeta'_l$. Hence, we cannot show approximation properties according to the concave corners because it is not clear which of the exact solutions will be approximated. For this reason we will back down on the given measurements for the following approximation result.

Theorem 8.2.23:

Let $\{\widetilde{P}_j\}_{j=1}^{M'}$ be a point set obtained by Algorithm 8.2.22. If $\|C - x_E(h_k)\|_2 + \|C - x_R(h_k)\|_2 < ct_k$ (diffraction) holds, then

$$\begin{aligned} \forall j \leq M' : \|\widetilde{P}_j - x_E(h_k)\|_2 + \|\widetilde{P}_j - x_R(h_k)\|_2 &\leq ct_k + 2\widetilde{\varepsilon} \\ \exists j \leq M' : \|\widetilde{P}_j - x_E(h_k)\|_2 + \|\widetilde{P}_j - x_R(h_k)\|_2 &\geq ct_k - 2\widetilde{\varepsilon} \end{aligned}$$

for all $k \leq N$. If $\|C - x_E(h_k)\|_2 + \|C - x_R(h_k)\|_2 > ct_k$ (reflection) then

$$\begin{aligned} \forall j \leq M' : \|\widetilde{P}_j - x_E(h_k)\|_2 + \|\widetilde{P}_j - x_R(h_k)\|_2 &\geq ct_k - 2\widetilde{\varepsilon} \\ \exists j \leq M' : \|\widetilde{P}_j - x_E(h_k)\|_2 + \|\widetilde{P}_j - x_R(h_k)\|_2 &\leq ct_k + 2\widetilde{\varepsilon} \end{aligned}$$

for all $k \leq N$. This means, the measurements are approximated with an accuracy of $2\widetilde{\varepsilon}$.

Proof:

We first show that the algorithm reconstructs a concave corner and hence is well-defined. Step 1 was discussed in Section 8.1. According to Theorem 8.2.21 steps 2-4 return a complete ε -separation $\{\widehat{\zeta}_j\}_{j=1}^M$ that is adapted to the measurement. Let $\{\zeta_j\}_{j=1}^{M'}$ be a set of directions such that an exact solution for those directions exists, i.e. there exist $r \in \mathbb{R}^{M'}$ and a set of points $\{P_j\}_{j=1}^{M'}$ with $P_j = C + r_j \zeta_j$ for $j = 1, \dots, M'$ such that

$$\begin{aligned} \min_{j \leq M'} \|P_j - x_E(h_k)\|_2 + \|P_j - x_R(h_k)\|_2 &= ct_k \text{ for reflection,} \\ \max_{j \leq M'} \|P_j - x_E(h_k)\|_2 + \|P_j - x_R(h_k)\|_2 &= ct_k \text{ for diffraction,} \end{aligned}$$

for all $k \leq N$. Since $\{\widehat{\zeta}_j\}_{j=1}^M$ is complete, there exists a subset $\{\widehat{\zeta}_j\}_{j \in J}$ such that for all $j' \leq M'$ there exists a $j \in J$ with $\zeta_{j'} \in \widehat{\zeta}_j$. Hence, using Theorem 8.2.19, we know that \widetilde{Q} has a concave corner \widetilde{r} with $\widetilde{r} - \widetilde{b}_k \leq \widetilde{\varepsilon}$ for all $k = 1, \dots, N$. Thus Step 6-8 will return a concave corner and moreover the algorithm will always terminate.

Now let us show that the returned set $\{\widetilde{P}_j\}_{j=1}^{M'}$ has the claimed approximation properties. We will first consider the case where $\|C - x_E(h_k)\|_2 + \|C - x_R(h_k)\|_2 > ct_k$ (reflection), the second case then follows analogously. Therefore, let $b_{kj} = \min\{r > 0 \mid C + r\zeta_j \in E_{t_k, h_k}\}$ or $b_{kj} = \infty$ if there does not exist such an $r > 0$. We first consider the case $b_{kj} = \infty$. It follows that $\widetilde{b}_{kj} = \infty$ because $\{\widehat{\zeta}_j\}_{j=1}^M$ is adapted to the measurement. Hence $\|\widetilde{P}_j - x_E(h_k)\|_2 + \|\widetilde{P}_j - x_R(h_k)\|_2 > ct_k > ct_k - 2\widetilde{\varepsilon}$.

Now consider $b_{kj} < \infty$, i.e. $C + b_{kj}\zeta_j \in E_{t_k, h_k}$ by definition and hence

$$\|C + b_{kj}\zeta_j - x_E(h_k)\|_2 + \|C + b_{kj}\zeta_j - x_R(h_k)\|_2 = ct_k.$$

It follows that $\widetilde{b}_{kj} < \infty$ by definition. For $\widetilde{r}_j < b_{kj}$ it follows that $\|\widetilde{P}_j - x_E(h_k)\|_2 + \|\widetilde{P}_j - x_R(h_k)\|_2 > ct_k > ct_k - 2\widetilde{\varepsilon}$ because b_{kj} is by definition the smallest real value such that $C + b_{kj}\zeta_j \in E_{t_k, h_k}$, otherwise we obtain

$$\begin{aligned} \|\widetilde{P}_j - x_E(h_k)\|_2 + \|\widetilde{P}_j - x_R(h_k)\|_2 &= \|C + (\widetilde{r}_j + b_{kj} - b_{kj})\zeta_j - x_E(h_k)\|_2 + \\ &\quad + \|C + (\widetilde{r}_j + b_{kj} - b_{kj})\zeta_j - x_R(h_k)\|_2 \\ &\geq \|C + b_{kj}\zeta_j - x_E(h_k)\|_2 - 2|\widetilde{r}_j - b_{kj}| \\ &\quad + \|C + b_{kj}\zeta_j - x_R(h_k)\|_2 \\ &= ct_k - 2(\widetilde{r}_j - b_{kj}) \\ &\geq ct_k - 2(\widetilde{r}_j - \min_{k \leq N} \widetilde{b}_{kj}) \\ &\geq ct_k - 2\widetilde{\varepsilon}. \end{aligned}$$

Now because $\tilde{r} \notin \tilde{Q}$ for each $k \leq N$ there exists a $j_k \leq M'$ such that $\tilde{r}_{j_k} \geq \tilde{b}_{kj_k}$. Thus it follows

$$\begin{aligned} \|\tilde{P}_{j_k} - x_E(h_k)\|_2 + \|\tilde{P}_{j_k} - x_R(h_k)\|_2 &= \|C + (\tilde{r}_{j_k} + b_{kj_k} - b_{kj_k})\tilde{\zeta}_{j_k} - x_E(h_k)\|_2 + \\ &\quad + \|C + (\tilde{r}_{j_k} + b_{kj_k} - b_{kj_k})\tilde{\zeta}_{j_k} - x_R(h_k)\|_2 \\ &\leq \|C + b_{kj_k}\tilde{\zeta}_{j_k} - x_E(h_k)\|_2 + 2|\tilde{r}_{j_k} - b_{kj_k}| \\ &\quad + \|C + b_{kj_k}\tilde{\zeta}_{j_k} - x_R(h_k)\|_2 \\ &= ct_k + 2|\tilde{r}_{j_k} - b_{kj_k}| \end{aligned}$$

First consider the case $\tilde{r}_{j_k} \geq b_{kj_k}$, then

$$|\tilde{r}_{j_k} - b_{kj_k}| = \tilde{r}_{j_k} - b_{kj_k} \leq \tilde{r}_{j_k} - \tilde{b}_{kj_k} \leq \tilde{r}_{j_k} - \min_k \tilde{b}_{kj_k} \leq \tilde{\varepsilon}.$$

Now let $\tilde{r}_{j_k} < b_{kj_k}$, then

$$|\tilde{r}_{j_k} - b_{kj_k}| = b_{kj_k} - \tilde{r}_{j_k} \leq b_{kj_k} - \tilde{b}_{kj_k} \leq \tilde{\varepsilon}$$

and hence the statement follows.

Now let $\|C - x_E(h_k)\|_2 + \|C - x_R(h_k)\|_2 < ct_k$ (diffraction) hold. We again define $b_{kj} = \min\{r > 0 \mid C + r\tilde{\zeta}_j \in E_{t_k, h_k}\}$ (here this is always well defined and we need not to consider the case where $b_{kj} = \infty$). Analogously as for $\tilde{r}_j < b_{kj}$ it follows that $\|\tilde{P}_j - x_E(h_k)\|_2 + \|\tilde{P}_j - x_R(h_k)\|_2 < ct_k < ct_k + 2\tilde{\varepsilon}$ because b_{kj} is by definition the smallest real value such that $C + b_{kj}\tilde{\zeta}_j \in E_{t_k, h_k}$, otherwise we obtain as above

$$\begin{aligned} \|\tilde{P}_j - x_E(h_k)\|_2 + \|\tilde{P}_j - x_R(h_k)\|_2 &\leq ct_k + 2|\tilde{r}_j - b_{kj}| \\ &= ct_k + 2(\tilde{r}_j - b_{kj}) \\ &\leq ct_k + 2(\tilde{r}_j - \min_k b_{kj}) \\ &\leq ct_k + 2\tilde{\varepsilon}. \end{aligned}$$

Furthermore, there exists $j_k \leq M'$ such that $\tilde{r}_{j_k} \geq \tilde{b}_{kj_k}$ and hence

$$\|\tilde{P}_{j_k} - x_E(h_k)\|_2 + \|\tilde{P}_{j_k} - x_R(h_k)\|_2 \geq ct_k - 2|\tilde{r}_{j_k} - b_{kj_k}| \geq ct_k - 2\tilde{\varepsilon}.$$

□

8.3 Summary

Let us summarize the results obtained in this chapter. In the third step of our method we had to reconstruct the defect boundary ∂D for given measurements $\{\mathbf{t}_k, h_k\}_{k=1}^M$ such that

$$\int_{\partial D \cap E_{h_k, \mathbf{t}_k}} a(\mathbf{t}_k, y, h_k) dy \neq 0$$

holds. To construct a discrete version of this problem we defined a representation $\{P_j\}_{j=1}^M \subset \partial D$ of ∂D and assumed that ∂D can be approximated by the polygonal line ($n = 2$) or the triangulation ($n = 3$) $\Delta(P_1, \dots, P_M)$. In Theorem 8.0.2 it was shown that this assumption leads to a good approximation of defect boundaries ∂D that can be described by a Lipschitz-continuous function. Considering only star-shaped defects we introduced the center point C , the directions $\{\zeta_j\}_{j=1}^M \subset \mathcal{S}$ and the radii $\{r_j\}_{j=1}^M$ with $P_j = C + r_j \zeta_j$. Hence, to find an approximation of the defect boundary ∂D , we have to reconstruct C , $\{\zeta_j\}_{j=1}^M$ and $\{r_j\}_{j=1}^M$. We first analysed the question how to find a center point C in Section 8.1. Since the set of possible center points \mathfrak{C} highly depends on the defect boundary, the reconstruction of C is a complicated problem. However, we presented heuristic methods using a-priori information about the measurement set-up that provide good approximations for C in numerical experiments. Moreover, according to Theorem 8.1.1 and Corollary 8.1.2 it is suitable to assume that the reconstructed point $C \in \mathfrak{C}$ is an exact center point.

Next in Section 8.2 we discussed the reconstruction of $\{\zeta_j\}_{j=1}^M$ and $\{r_j\}_{j=1}^M$ for a given center point C . Therefore, the reconstruction problem was reformulated as a non-convex minimization problem w.r.t. $\{r_j\}_{j=1}^M$ with parameters $\{\zeta_j\}_{j=1}^M$. The spiderweb algorithm 8.2.8 was introduced that is able to find a local solution of this minimization problem. Since the directions $\{\zeta_j\}_{j=1}^M$ are parameters that have to be chosen, we first analysed the case where $\{\zeta_j\}_{j=1}^M$ were given exactly. This analysis was expanded to the case where the directions were only approximatively given and then further expanded to the situation of unknown directions $\{\zeta_j\}_{j=1}^M$. We provided approximation results for all three cases and finally represented Algorithm 8.2.22 which can be used to reconstruct $P_j = C + r_j \zeta_j$ for $j = 1, \dots, M$ with an approximation accuracy of $2\tilde{\varepsilon}$ with respect to the measurements $\{\mathbf{t}_k, h_k\}$ (see Theorem 8.2.23). Figure 8.12 illustrates the reconstruction of a defect boundary with two different ε -separations. For further and more detailed numerical examples we refer to the next chapter.

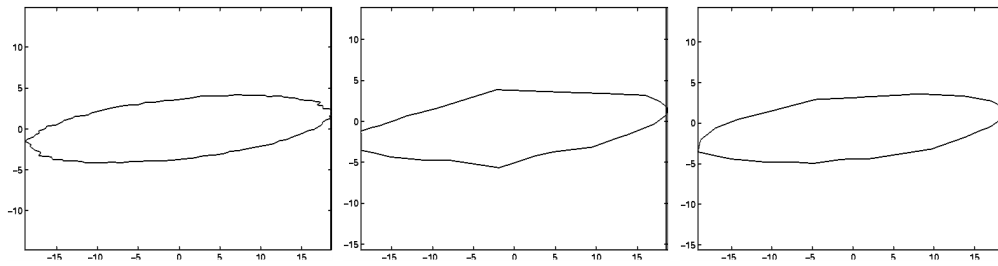


Fig. 8.12: Reconstruction of the defect boundary using Algorithm 8.2.22. Original boundary (left) and reconstruction using an ε -separation with $M = 500$ (middle) and $M = 1200$ (right) sets.

9 Numerical Results

After having introduced our inversion method in the last three chapters, we want to analyse some numerical results for the presented methods. In Chapter 6 we have discussed some theoretical approximation results for MP. This will now be confirmed with some numerical examples in the first section. Furthermore, the parameter estimation with the Newton-HOMP algorithm 6.3.1 will be discussed here. In the later sections of this chapter we will also discuss the other steps of the inversion method and finally give some reconstruction examples where all three steps are applied consecutively.

9.1 Approximation Results for OMP and Newton-HOMP

Analysing the deconvolution results of OMP and Newton-HOMP using simulated or real NDT-data is a complicated task. Although parameters of the transmitted ultrasound impulse might be known, these parameters might change due to effects in the steel component. Moreover, most signals like the lateral wave and the back wall echo are not the result of a single reflected or diffracted impulse but consist of several strongly overlapping impulses. The deconvolution method will therefore not reconstruct each impulse echo separately but approximate the complete lateral wave or back wall echo with one impulse that is most likely somewhere “in the middle” of the original impulses. Compared with the ideal \mathbf{u} of the form

$$\mathbf{u} = f * X$$

where f is a sampled ultrasound impulse and X is a vector obtained by e.g. our geometrical model, this means that the sparse deconvolution methods applied to NDT-data will not reconstruct f and X but find vectors \tilde{f} and \tilde{X} where $\tilde{f} * \tilde{X} \approx f * X$. Hence, analysing the approximation results for OMP or Newton-HOMP on NDT-data is futile as we know beforehand that the reconstruction will differ from the original data. However, in this section we show three examples where the overlapping effect does not occur and thus we can test the approximation properties of the deconvolution methods on this data (see also the numerical examples in [80]).

Let us first consider an example signal consisting of 4 ultrasound impulses where one impulse approaches towards one of the three fixed impulses, i.e. a function $X(t)$ in time that is only non-zero at $t = 0.5, 3.4, x, 7.2$ and we choose $x = 6.96, 7.07, 7.15, 7.18, 7.19$. Hence, the third impulse is approaching towards the fourth one and the correlation

between both will get stronger the larger x is chosen. The function X is then convolved with an ultrasound impulse that has a bandwidth of 50Mhz^2 , a frequency of 4Mhz and a shift of 0.52 . The result is shown in Figure 9.1. Then we use OMP to reconstruct the times and amplitudes, i.e. we use HOMP with $\varepsilon_0 = 0$, $t_k = t_k = 0.01\mu s$, $w = 1$, $L = 5$, $\varepsilon_1 = 0$ and $\varepsilon_2 = 1.25/\|u\|_2$. The original and reconstructed amplitudes and times are shown in Table 9.1. As one can see, OMP is able to reconstruct times and amplitudes of the two non-interfering impulses. Moreover, it can even distinguish the interfering signals upto a time difference of $13\mu s$. However, if x comes too close to 7.2 , OMP can only reconstruct one impulse that approximates both original impulses. The reconstructed amplitude then gets close to the sum of both amplitudes.

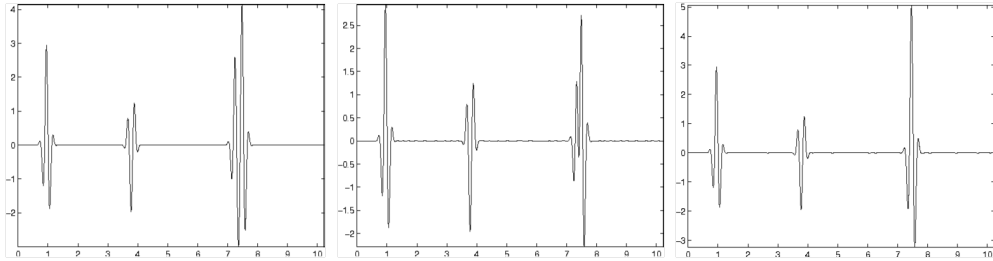


Fig. 9.1: Function u with four ultrasound impulses at time $0.5, 3.4, x, 7.2$ where $x = 6.96$ (left), $x = 7.07$ (middle) and $x = 7.15$ (right); time in microseconds.

original signal	arrival times (μs)				amplitudes			
	0.50	3.40	x	7.20	3.0	-2.0	2.5	4.0
time x	arrival times (μs) obtained by OMP				amplitudes obtained by OMP			
6.96	0.50	3.40	6.96	7.20	3.0028	-1.9937	2.5001	3.9988
7.07	0.50	3.40	6.95	7.21	3.0023	-2.0012	-1.3483	2.4355
7.15	0.50	3.40	7.18		2.9950	-1.9958	5.1517	
7.18	0.50	3.40	7.19		3.0030	-1.9950	6.2750	
7.19	0.50	3.40	7.20		2.9996	-2.0000	6.3707	

Table 9.1: Parameter estimation results for four interfering echoes when one arrival time gets close to another.

In a second example we want to analyse the behaviour of OMP on noisy data. Again, we consider a peak function X with peaks at time $t = 0.5, 3.4, 3.75, 7.2$ and corresponding amplitudes $3.0, -2.0, 4.0, 2.5$. Now, before X is convolved with an ultrasound impulse with bandwidth 20Mhz^2 , frequency of 8Mhz and a shift of 0.52 , we add Gaussian noise to the data. Using the OMP with settings as in the last experiment we reconstruct the original peaks for different noise levels. Figure 9.2 shows the original signal without noise and two signals that are obtained using different noise levels. Table 9.2 lists the reconstructed times and amplitudes. For a better comparison we have also evaluated the SNR of the noisy convolved data and of the reconstructed peak signal convolved again

with the impulse f . The data shows that while the amplitudes are influenced by the noise level, the times are nearly perfectly reconstructed even for high noise levels. This confirms our assumption that our reconstruction method should base upon the peaks' positions and neglect the amplitude itself.

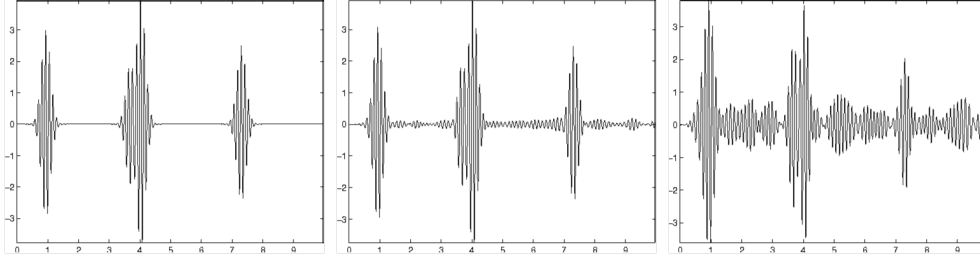


Fig. 9.2: Obtained signal $u = f * X$ for different noise levels on X : without noise (left), with SNR of 19.34 (middle) and with SNR of 4.75 (right); time in microseconds.

original signal	arrival times (μs)				amplitudes				
	0.50	3.40	3.75	7.20	3.0	-2.0	4.0	2.5	
SNR	arrival times (μs) obtained by OMP				amplitudes obtained by OMP				estimation SNR
19.34	0.50	3.40	3.75	7.20	3.1000	-1.9727	3.9863	2.4562	34.22
13.10	0.50	3.40	3.75	7.20	3.1071	-1.9352	3.9093	2.4312	30.88
8.15	0.50	3.40	3.75	7.20	3.1085	-1.7118	3.9806	2.4363	25.37
4.75	0.50	3.40	3.75	7.19	4.0206	-2.5048	3.4768	1.9641	10.33

Table 9.2: Parameter estimation results for four echoes with different SNR with OMP.

In a last example we want to analyse the use of the Newton-HOMP technique to reconstruct the ultrasound impulse parameters. Therefore we consider a peak function X with peaks at $t = 0.1, 0.8, 1.8$ microseconds and corresponding amplitudes $1, 0.4, -0.7$. We have added Gaussian noise with different noise levels to X before convolving with the ultrasound impulses. The obtained data is used to reconstruct the impulse parameters with the Newton-HOMP method where we set $w = 1$, $\varepsilon_0 = \infty$ (i.e. we use MP), $\varepsilon_1 = \varepsilon_2 = 0$ and $L = 3$ (i.e. we reconstruct exactly three peaks). For a better accuracy of the method we analysed three different time samplings, i.e. $t_{k+1} - t_k = 0.01, 0.005, 0.0033$ microseconds. Figure 9.3 illustrates three examples of the obtained data $u = f * X$. Tables 9.3, 9.4 and 9.5 show the starting parameter obtained as described in subsection 6.2.1 and the parameters obtained with Newton-HOMP for different discretizations t_k where each table presents the data for a fixed ultrasound impulse. As one can see, the approximation accuracy seems to dependent highly on the impulse parameters while the discretization distance has no big influence on the results. The starting guess is quite accurate for all three impulses and for low noise levels Newton-HOMP is often not able to improve the results. For higher noise levels the method is able to reconstruct a good parameter estimation. Nevertheless, while the frequency is always approximated quite

well, the algorithm may fail to reconstruct bandwidth and shift if the noise level is too high. However, note that although the parameters might differ a lot, the reconstructed impulse is quite similar to the original one as illustrated in Figure 9.4 (see also Table 9.4, starting guess for highest noise level). We will use the Newton-HOMP method with the same parameter set-up and a discretization distance of $t_{k+1} - t_k = 0.005$ in the next sections to approximate the unknown impulse function parameters. Note that this only needs to be done once for each testing set-up. The obtained parameters can then be used to deconvolve the data using the HOMP algorithm only. Hence we do not count the runtime of Newton-HOMP to the overall runtime of our reconstruction method. Nevertheless, each result shown in Table 9.3, 9.4 and 9.5 was obtained in less than 3, 6 or 10 seconds depending on the chosen discretization distance 0.01, 0.005 or 0.0033 and thus the method is fast enough to be used as a calibration before the real testing.

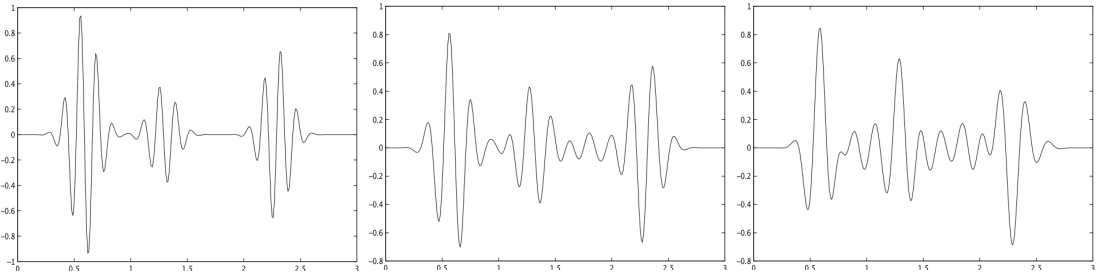


Fig. 9.3: Obtained signal $u = f * X$ for different noise levels on X and different impulse parameters: bandwidth 40MHz^2 , frequency 7MHz , shift $\pi/2$ without noise (left), bandwidth 30MHz^2 , frequency 5MHz , shift $\pi/4$, SNR of 13.31 (middle), bandwidth 50MHz^2 , frequency 4MHz , no shift, SNR of 8.55 (right).

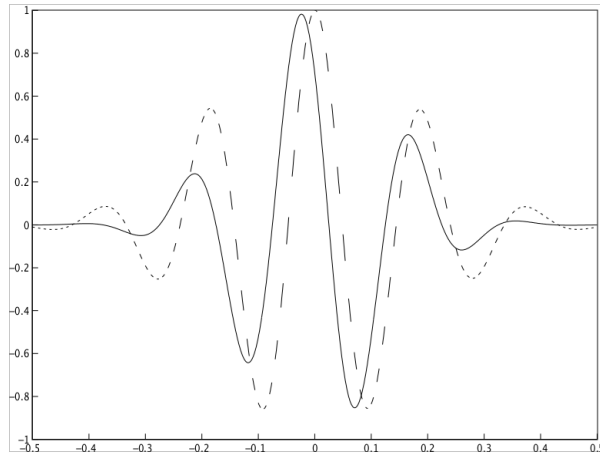


Fig. 9.4: Two impulse functions with parameters: bandwidth 30MHz^2 , frequency 5MHz and shift $\pi/4$ (solid); bandwidth 17.3008, frequency 5.2236 and shift 3.1175 (dashed).

original signal	bandwidth ρ Mhz ²		frequency ψ Mhz	shift ϕ			
	40		7	$\pi/2$			
SNR	starting guess			final estimate			discretization
	ρ	ψ	ϕ	ρ	ψ	ϕ	$t_k - t_{k+1}$
no noise	40	7	1.6346	40.0002	7	1.5722	0.01
				40.0002	7	1.5722	0.005
				40.0002	7	1.5722	0.0033
19.21	38.0837	6.9432	1.1849	36.1429	6.9879	1.1722	0.01
				36.1429	6.9879	1.1722	0.005
				36.1429	6.9879	1.1722	0.0033
13.81	37.1189	6.9597	1.2305	34.5467	6.9725	1.1207	0.01
				34.4296	6.9592	1.1554	0.005
				33.5692	6.9784	1.2744	0.0033
5.86	67.962	6.7924	1.546	40.1333	7.0088	1.5912	0.01
				40.0346	6.9671	1.6359	0.005
				39.933	6.9575	1.5156	0.0033

Table 9.3: Impulse parameter estimation results with different SNR with the Newton-HOMP method.

original signal	bandwidth ρ Mhz ²		frequency ψ Mhz	shift ϕ			
	30		5	$\pi/4$			
SNR	starting guess			final estimate			discretization
	ρ	ψ	ϕ	ρ	ψ	ϕ	$t_k - t_{k+1}$
no noise	30.0003	5	0.8164	29.9982	5	0.7863	0.01
				29.9982	5	0.7863	0.005
				29.9982	5	0.7863	0.0033
21.58	27.698	5.0756	1.03	29.5583	5.0238	1.1136	0.01
				29.7268	5.0259	0.9617	0.005
				29.6801	5.0254	1.0118	0.0033
13.31	28.8636	5.119	0.7114	30.1959	5.107	0.6232	0.01
				30.4795	5.1128	0.7245	0.005
				30.3882	5.1113	0.6907	0.0033
5.62	17.3008	5.2236	3.1175	19.018	5.1615	0.0529	0.01
				19.8805	5.1661	0.0045	0.005
				19.4016	5.1631	3.1252	0.0033

Table 9.4: Impulse parameter estimation results with different SNR with the Newton-HOMP method.

original signal	bandwidth ρ Mhz ²	frequency ψ Mhz	shift ϕ				
	50	4	0				
SNR	starting guess			final estimate			discretization
	ρ	ψ	ϕ	ρ	ψ	ϕ	$t_k - t_{k+1}$
no noise	50	4	0.0095	50	4	0.0002	0.01
				50	4	0.0002	0.005
				50	4	0.0002	0.0033
20.24	52.6904	4.1412	0.0724	51.5772	4.0499	0.0036	0.01
				51.2958	4.0462	0.0933	0.005
				51.4034	4.048	0.0634	0.0033
13.74	56.4492	4.0348	0.1026	58.2268	3.9613	0.1095	0.01
				58.1558	3.9617	0.0529	0.005
				58.179	3.9592	0.116	0.0033
8.55	52.86	4.1184	2.9861	56.0173	4.1511	2.9556	0.01
				55.8008	4.1537	2.9953	0.005
				55.8705	4.1533	2.982	0.0033

Table 9.5: Impulse parameter estimation results with different SNR with the Newton-HOMP method.

9.2 Clustering Results for DBSCAN and DBCLAN

Now let us analyse the second step of our inversion method. Therefore, we first analyse the adapted DBSCAN algorithm 7.2.1 that can be used for non-overlapping clusters. We use the parameters $P_{\min} = 2$ and $\gamma = c^2 = 5.92^2$ for an equal treatment of time and space variables. As input data we use a sparse logical matrix containing four clusters simulating a deconvolved ToFD image with lateral wave, back wall echo, a cluster typical for pore defects and a cluster representing a crack defect. Furthermore, we randomly set elements of the data to 1 to simulate noise. Figure 9.5 illustrates the input data without noise, 0.1% noised, 0.5% noised and 1% noised. We reconstruct the different clusters using the DBSCAN algorithm with $\varepsilon = 0.16$ for the highest noise level and $\varepsilon = 0.26$ otherwise. The different reconstructed clusters are shown in Figures 9.6 to 9.9 where we have neglected minor clusters with only a few points. We observe that with increasing noise level also the clusters are corrupted by noise, Table 9.6 lists the number of elements obtained in the reconstructed clusters. However, the algorithm is still able to separate the different clusters.

Now we analyse our clustering algorithm DBCLAN 7.3.1 for overlapping clusters. Therefore we simulate deconvolved data as a sparse matrix that contains three clusters representing the lateral wave, the back wall echo and a pore defect. Here, the pore defect signal and the back wall echo intersect and hence cannot be separated by clustering algorithms that only use spatial information. Again we consider different noise levels where 0.1%, 0.5% and 1% of the elements are set to 1. The input data is illustrated in

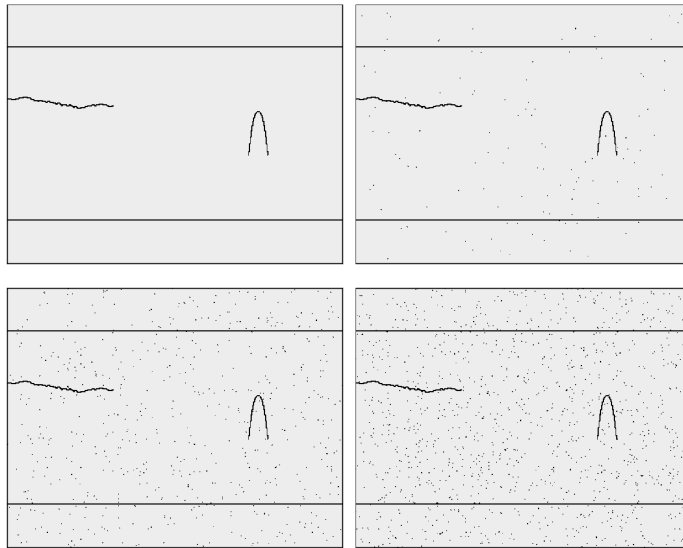


Fig. 9.5: Sparse input data with 4 clusters and different noise levels; lrtb: no noise, 0.1% noise, 0.5% noise and 1% noise.



Fig. 9.6: Reconstructed lateral wave cluster for different noise levels.

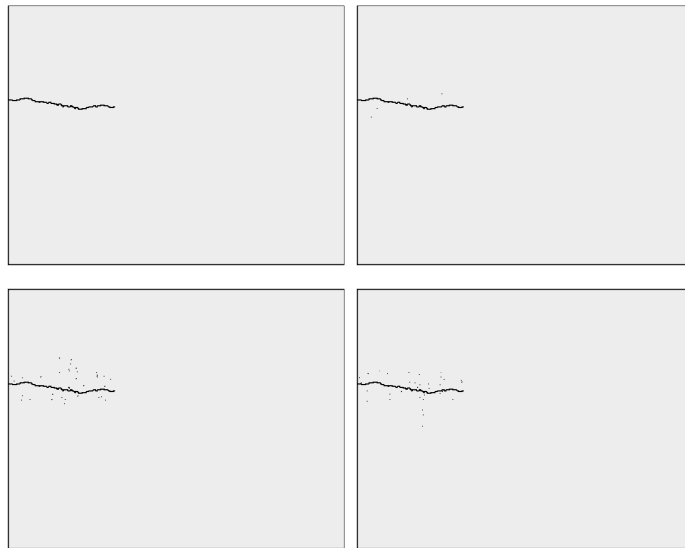


Fig. 9.7: Reconstructed crack cluster for different noise levels.

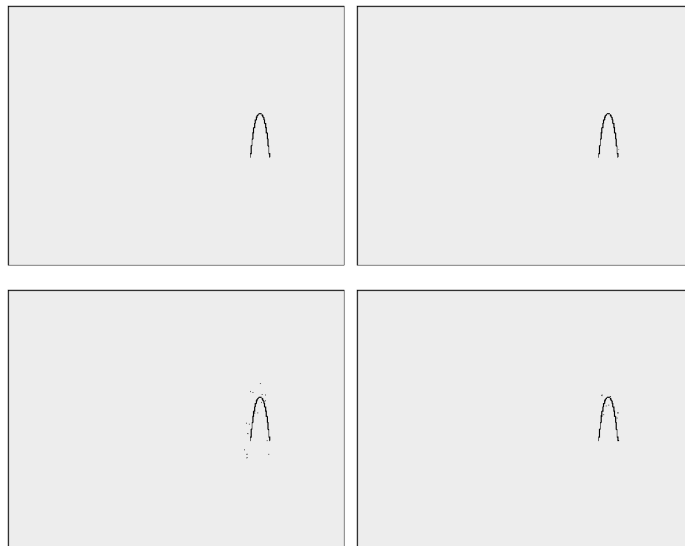


Fig. 9.8: Reconstructed pore cluster for different noise levels.



Fig. 9.9: Reconstructed back wall cluster for different noise levels.

noise	lateral wave cluster	crack cluster	pore cluster	back wall cluster
0	400	167	122	400
0.1%	418	171	123	426
0.5%	522	201	139	539
1%	536	200	132	532

Table 9.6: Number of elements in the reconstructed clusters.

Figure 9.10. Since for core points with more than two neighbours one or more duplicates are created during the algorithm, it is crucial to keep the number of neighbours small. Thus, we use the 8-neighbourhood as already shown in Chapter 7. The results obtained with different noise levels are illustrated in Figures 9.11 to 9.13, the number of points in each cluster is shown in Table 9.7. We note that the algorithm is able to separate all clusters even for strong noise. Furthermore, the number of noise peaks added to each cluster is very small. Nevertheless, for high noise levels the algorithm becomes unstable in the sense that it may truncate some clusters as it can be seen in Figure 9.12.

noise	lateral wave cluster	pore cluster	back wall cluster
0	400	316	404
0.1%	401	316	407
0.5%	406	308	409
1%	416	323	404

Table 9.7: Number of elements in the reconstructed clusters.

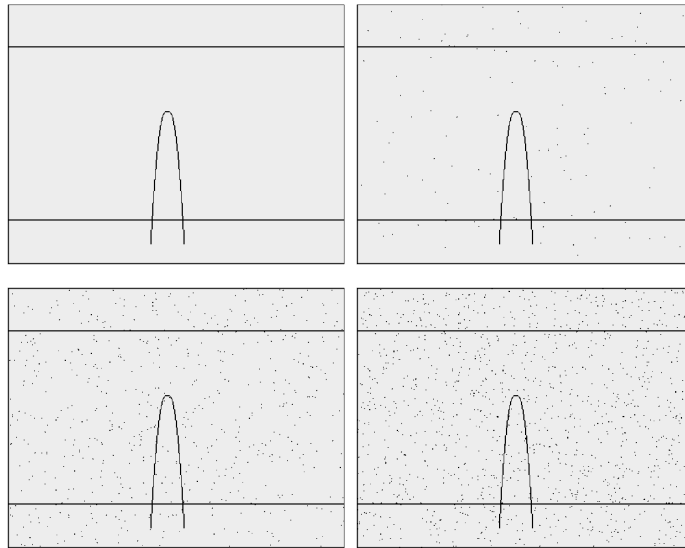


Fig. 9.10: Sparse input data with 3 clusters and different noise levels; lrtb: no noise, 0.1% noise, 0.5% noise and 1% noise.



Fig. 9.11: Reconstructed lateral wave cluster for different noise levels.

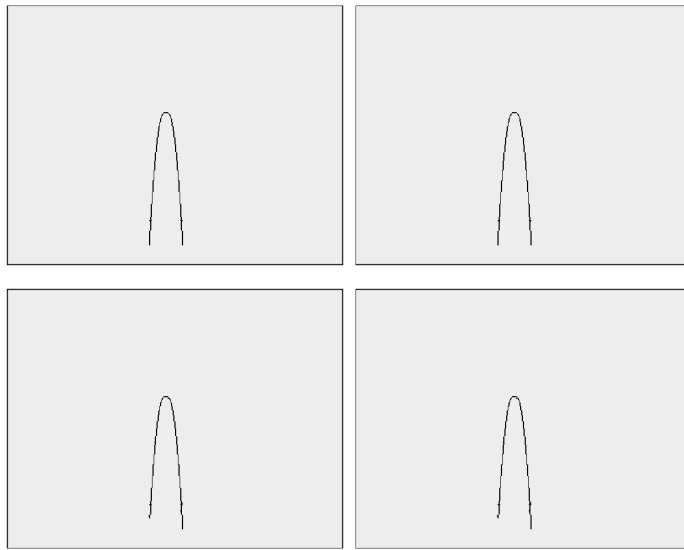


Fig. 9.12: Reconstructed pore cluster for different noise levels.

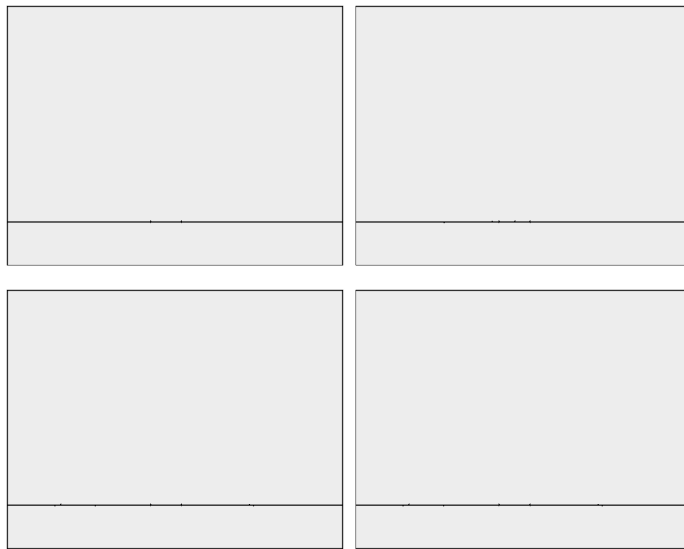


Fig. 9.13: Reconstructed back wall cluster for different noise levels.

9.3 Reconstruction with the Spiderweb Algorithm

The approximation properties of the third reconstruction step have already been theoretically analysed in Chapter 8. Nevertheless, to get an impression of its approximation quality, we are going to discuss two examples in this section. In the first example we analyse the approximation results of the spiderweb algorithm with an ε -separation for decreasing ε . Therefore, consider the two-dimensional case with a circle defect at position $(0, 7)$ with radius 1. For the positions $(x, 0)$ with $x = -50, -49.5, \dots, 50$ let the distance to the nearest and the most distant point on the defect be given. Figure 9.14 shows the distances plotted against x . These points form a cluster that can be seen as a cluster representing a pore defect in a ToFD set-up containing (nearly) perfect information. We assume that the center point $(0, 7)$ is exactly given and use the spiderweb algorithm to reconstruct the defect using different ε . Figure 9.15 illustrates the obtained results. As one can see, the reconstructions are getting better the smaller we choose ε , therefore see also Table 9.8 where the maximal distance to the center point is listed. For a good reconstruction, this value should be close to the radius 1.

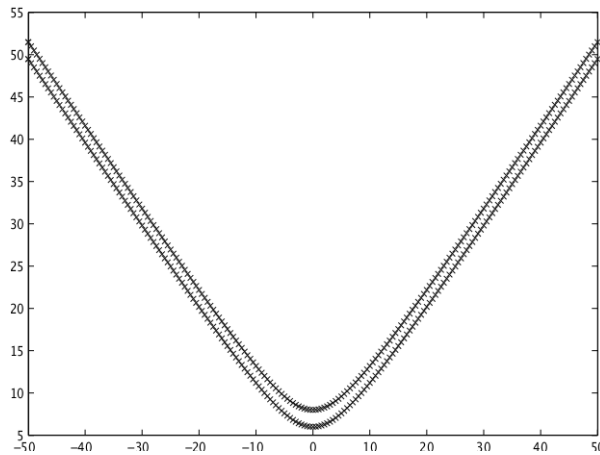


Fig. 9.14: Minimal and maximal distance to the circle defect for different positions x . These points form a cluster that is similar to clusters seen in ToFD set-ups for pore defects.

ε	0.08	0.05	0.01	0.005	0.001	0.0005
$\max_k \ C - P_k\ _2$	2.1122	1.3063	1.2370	1.2279	1.2090	1.1966

Table 9.8: Maximal distance between the center point and the reconstructed points.

In the second example we analyse the behaviour of the last reconstruction step when the input data is noisy, i.e. if the given minimal and maximal distance to the defect is noisy. This is similar to noisy NDT data where the arrival times of an ultrasound impulse cannot be determined exactly. We use the same defect as before since here the optimal center point $(0, 7)$ as well as a suitable error estimation using the maximal distance to

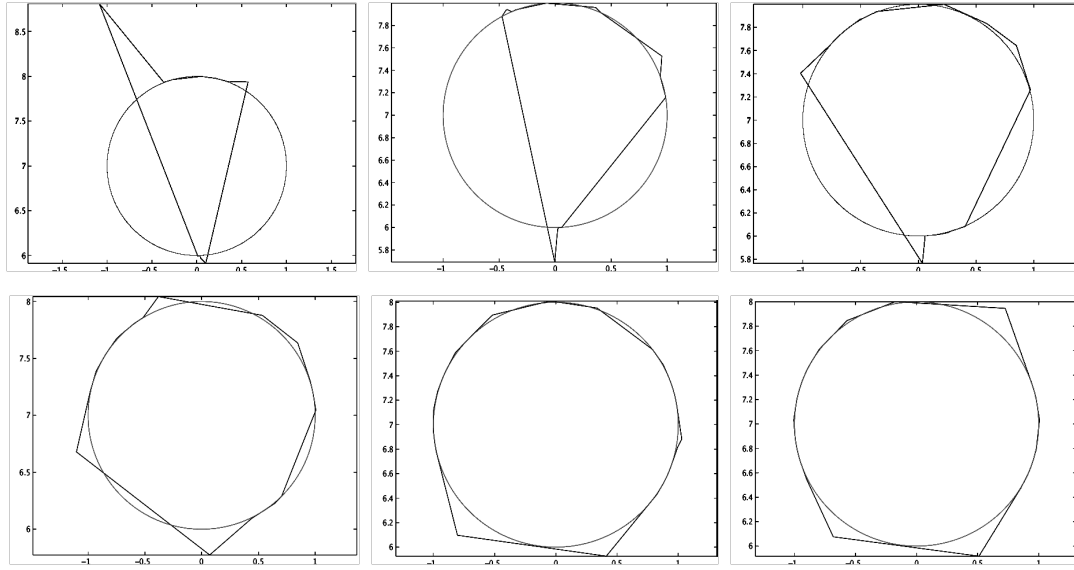


Fig. 9.15: Reconstructed defect and the original defect (circle) for different ε -separations; lrtb: $\varepsilon = 0.08, 0.05, 0.01, 0.005, 0.001, 0.0005$.

the point $(0, 7)$ is known. Furthermore, we also analyse an approximation method for the center point C that is based on the general heuristic approach presented in Section 8.1 where

$$C = \arg \max_{C' \in \mathbb{R}^n} R(C')I(C')$$

with

$$R(C') = \min_{(x,d) \in \text{Cl}} \left| \|C' - (x, 0)\|_2 - d \right|,$$

Cl is the cluster and I is a probability function. Given a cluster where the shortest and the longest distance to a defect is given, i.e. the arrival time for the first reflection and last diffraction signal is known, we can choose I in a clever way. We use the assumption that for two peaks $(x, d_1), (x, d_2)$ in the cluster with $d_1 < d_2$ the first peak (x, d_1) was obtained by a point $P_1 \in \partial D$ that is close to the probe position $(x, 0)$ while the second peak (x, d_2) was obtained by a point $P_2 \in \partial D$ more distant to $(x, 0)$. The center point C usually is somewhere “in the middle” of the defect and thus, the distance between the center point C and $(x, 0)$ lies in the interval (d_1, d_2) . Hence we choose

$$I(C') = \begin{cases} 1 & d_1 < \|C' - (x, 0)\|_2 < d_2 \text{ for all } (x, d_1), (x, d_2) \text{ in the cluster with } d_1 < d_2 \\ 0 & \text{otherwise} \end{cases}.$$

For computational reasons, we will not solve the minimax problem as suggested in Section 8.1 but just sample the functions R and I on a discrete grid. Figure 9.16 illustrates the

functions R , I and RI in the noiseless case, compare also 8.4. Table 9.9 shows the reconstructed center points for the cluster where noise is added to the distances. We consider uniformly distributed noise with maximum absolute value of 0, 0.1, 0.25 and 0.4. As one can see, the approximation of the center point is very stable.

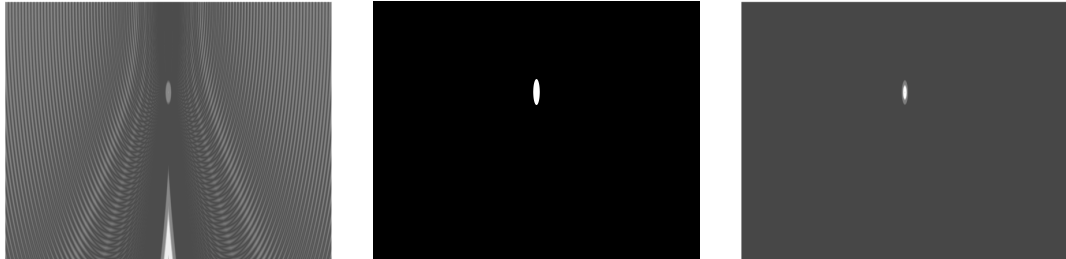


Fig. 9.16: Estimation of the center point C as the maximum of $R(C')I(C')$ (right) where R is a distance function (left) and I is a probability function (middle).

Noise	0	0.1	0.25	0.4
C	(0, 7)	(0, 7)	(-0.01, 7)	(0, 7.01)

Table 9.9: Reconstructed center points for different noise levels.

Now we use the spiderweb algorithm to approximate the defect where we choose an ε -separation with $\varepsilon = 0.001$. The obtained reconstruction results are shown in Figure 9.17. Table 9.10 lists the maximal distance to the original center point (0, 7). We note that the size of the reconstructed defect seems to increase with the noise. However, the approximation is still suitable and the method appears to be stable even for noisy data.

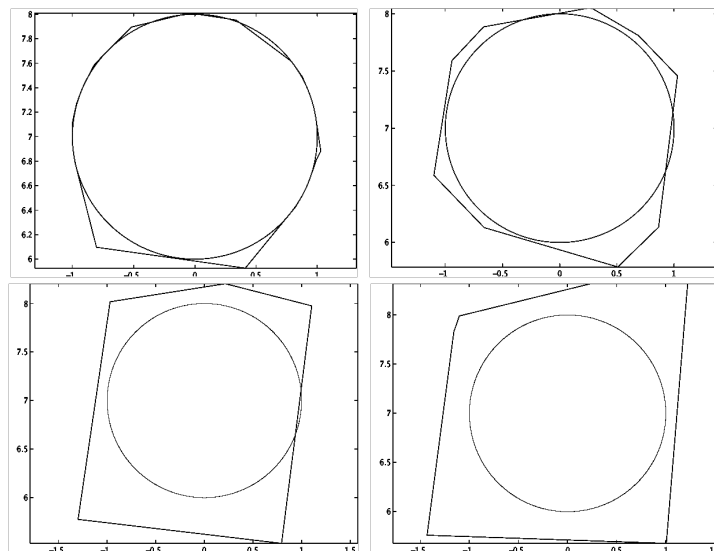


Fig. 9.17: Reconstructed defects for different noise levels.

Noise	0	0.1	0.25	0.4
$\max_k \ C - P_k\ _2$	1.209	1.3159	1.7851	1.8931

Table 9.10: Maximal distance between the center point and the reconstructed points for different noise levels.

9.4 Numerical Results for the Complete Inversion Method

In the last sections, we have analysed each reconstruction step on its own. Now let us consider some examples where we apply the complete reconstruction method to simulated and real data. Furthermore, remembering Chapter 3 we will use the SAFT method for the simulated data to compare the reconstruction results with our method.

9.4.1 Simulated ToFD Pore

Our first example is a ToFD measurement set-up. We simulate a 20mm thick component with a weld at position $z = 0$. The surface of the component is the plane $x = 0, y = 0$. A pore with a diameter of 3mm is at position (0mm, -12mm, 0mm) in the weld. The probes move with a distance of 40mm to the weld alongside it starting at $x = -75$ mm and stopping at $x = 75$ mm (see Figure 9.18). The emitter is focussing on a point approximately at 2/3 of the weld height with an angle of aperture of 12 degree, i.e. the amplitude transmitted into direction s is given by

$$\text{sinc}(2.8 \arccos(\langle (0, -\sin 20^\circ, \cos 20^\circ), s \rangle)).$$

The emitted ultrasound impulse has a bandwidth of 30Mhz^2 , a frequency of 5Mhz and a shift of $\pi/4$. A-scans were sampled every 0.5mm starting shortly before the lateral wave arrives at $13\mu\text{s}$ with a sampling discretization of $0.01\mu\text{s}$. We apply the constants $c = c_l = 5.92$, $c_t = 3.23$ and attenuation $\gamma \approx (\psi/50)^2$. We simulate the measurement data with Gaussian noise on the geometrical model with different variances 0, 0.005, 0.01, 0.015, 0.02, 0.025 and 0.035 of the maximal amplitude, i.e. the variance is given by e.g. $0.005\|u\|_\infty$ where u is the data computed without noise. The results are shown in Figure 9.19.

As a first reconstruction step we approximate the impulse function parameters with the Newton-HOMP method using the first 50 a-scans of the data and a time sampling of $t_{k+1} - t_k = 0.005$. The obtained results are shown in Table 9.11. With the reconstructed impulse parameters we deconvolve the data using the HOMP algorithm 6.1.1 with the sampling grid $t_k = t_k$ and the parameters $L = 5$, $\varepsilon_1 = 0.05$, $\varepsilon_2 = 0.15$. Furthermore, we decrease ε_0 with increasing noise level to counteract the noisy data with more orthogonalisation steps in HOMP, i.e. we set $\varepsilon_0 = 1, 0.5, 0.3, 0.25, 0.2, 0.15, 0$ corresponding to the noise levels 0, 0.005, 0.01, 0.015, 0.02, 0.025, 0.035. This parameter setting has proven to provide good and stable results in our numerical analysis. However, since the defect

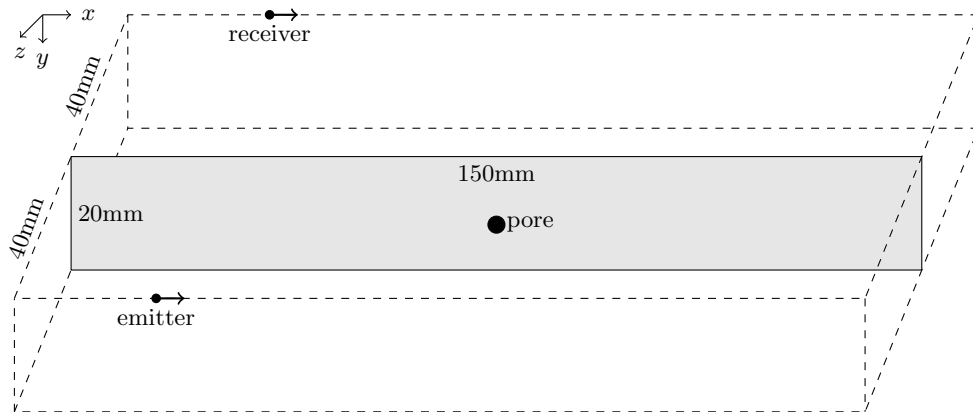


Fig. 9.18: Simulated testing set-up.

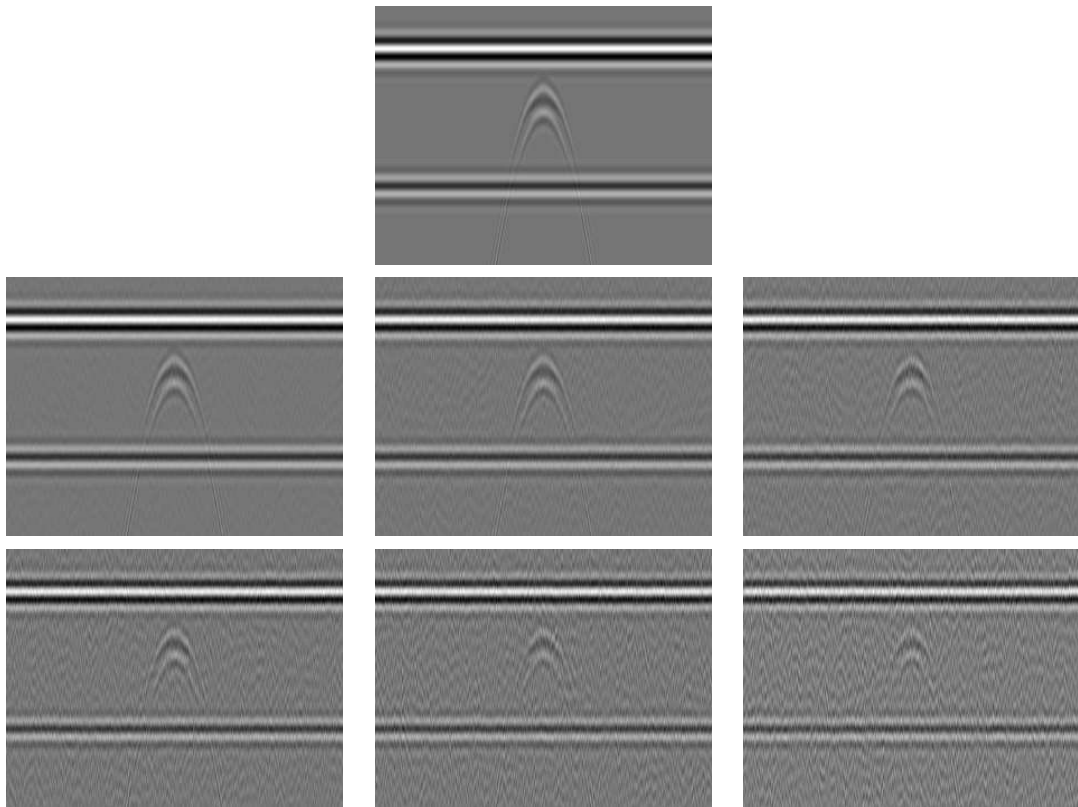


Fig. 9.19: Simulated ToFD measurement without noise (top) and with different Gaussian variance 0, 0.005, 0.01, 0.015, 0.02, 0.025, 0.035 multiplied with $\|\mathbf{u}\|_\infty$ where \mathbf{u} is the data computed without noise (top).

signals are close to the lateral wave, separating both signals in two different clusters in the second inversion step will be more complicated with increasing noise. In fact, the clustering will even fail for the two noisiest datasets if we do not apply special weights w_k in HOMP. Hence we set

$$w_k = e^{(\psi/50)^2 c t_k} \max(1 - 1.25|\langle f_k, f_L \rangle| - |\langle f_k, f_B \rangle|, 0) \quad (9.1)$$

where f_L is the sampled impulse function time shifted by $80/5.92$ being the arrival time of the lateral wave and f_B is time shifted by the arrival time of the back wall echo $2\sqrt{40^2 + 20^2}/5.92$. This means, the first factor of the weight w_k makes up for attenuation while the second part is small for functions f_k that highly correlate with the lateral wave or back wall echo. Since the lateral wave is usually a very strong signal we weight its correlation with a factor of 1.25. Applying the HOMP algorithm to the simulated data, where we choose w_k as above for the two highest noise levels and set $w_k = 1$ for all other levels, yields the data shown in Figure 9.20. As one can see, the defect as well as lateral wave and back wall echo are detected in the deconvolution process except for the weighted deconvolution where lateral wave and back wall echo are suppressed. However, with increasing noise also the number of peaks corresponding to noise is increasing. Hence, these peaks need to be sorted out later in the reconstruction process.

noise	ρ	ψ	ϕ
0	29.9962	5	0.8313
0.005	30.0913	5.0104	0.814
0.01	30.2287	4.982	0.9119
0.015	31.3854	4.9655	0.7546
0.02	30.5232	5.0072	0.6988
0.025	25.0391	4.9158	0.8354
0.035	27.906	4.9339	0.7747

Table 9.11: Impulse parameter estimation results for the modelled data shown in Figure 9.19.

After the first reconstruction step has been calculated, we cluster the data to separate the signal peaks caused by the defect from the lateral wave, the back wall echo and noise. We use our adaption of the DBSCAN algorithm 7.2.1 with the parameters $P_{\min} = 2$, $\gamma = 5.92^2$ and $\varepsilon = 1.54$. With increasing noise level the algorithm will also detect some minor clusters that only consist of noise but these clusters can be easily sorted out as the number of points in such a cluster is usually too small to correspond to a defect, i.e. these clusters normally contain less than 10 points. Moreover, we will detect one cluster for each lateral wave and back wall echo which we also neglect. The remaining main cluster contains the signal peaks caused by the modelled defect. Figure 9.21 shows the obtained cluster for each noise level. We note that even for strong noise the clustering algorithm is able to detect the significant defect peaks. Nevertheless, with increasing noise level the cluster does also contain more peaks that are caused by noise what will worsen the reconstruction.



Fig. 9.20: Non-zero elements of deconvolved data for different noise levels, special weights were used for the two highest noise levels (bottom, middle and right).

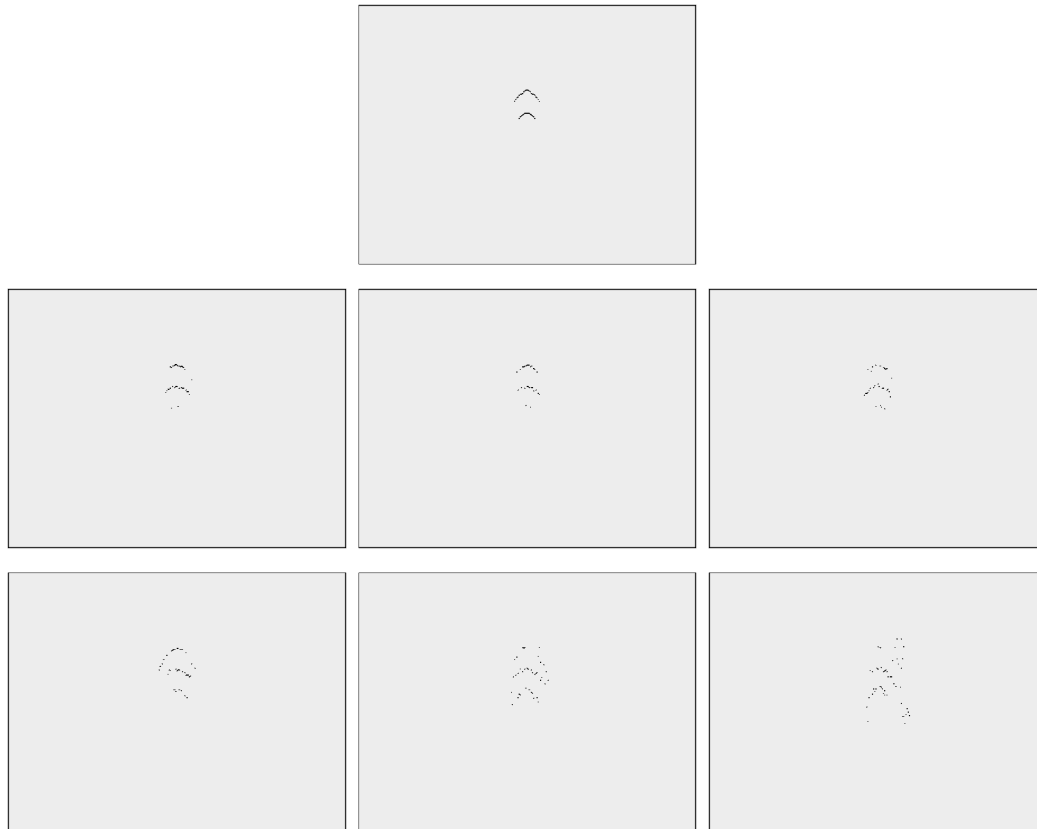


Fig. 9.21: Non-zero elements of the cluster containing the signal peaks of the modelled defect obtained with DBSCAN applied to the deconvolved data with different noise levels.

With the clusters calculated we now apply our last reconstruction step. To reconstruct the defect, we apply Algorithm 8.2.22. Since the given data is only two-dimensional and the inspection method aims to find defects in the weld plane, we will limit our reconstruction to the two-dimensional weld plane, i.e. to the plane $z = 0$. As a first step we need to find a center point C applying one of the heuristic methods presented in Section 8.1. Since the reconstructed clusters contain both reflection and diffraction signals we can use the same heuristic approach we discussed earlier in Section 9.3. Figure 9.22 illustrates the functions $R(C')$, $I(C')$ and $R(C')I(C')$ for the noiseless case. The obtained points C are shown in Table 9.12. Note that the distance between the reconstructed center points C and the original centre of the pore is always less than the pore radius and hence all reconstructed points are suitable.



Fig. 9.22: Estimation of the center point C as the maximum of $R(C')I(C')$ (right) where R is a distance function (left) and I is a probability function (middle).

Cluster	C	$\ C - (0, -12)\ _2$
no noise	(0,-12.1)	0.1
noise 0.005	(0,-10.6)	1.4
noise 0.01	(0.2,-10.7)	1.3153
noise 0.015	(0.1,-10.7)	1.3038
noise 0.02	(0.2,-12.2)	0.2828
noise 0.025	(-0.4,-12.1)	0.4123
noise 0.035	(0.5,-12.1)	0.5099

Table 9.12: Reconstructed center point and its distance to the original point.

Finally, we are able to reconstruct the defect using the spiderweb algorithm. For an ε -separation with $\varepsilon = 0.001$ we obtain the defect approximations shown in Figure 9.23 where the circle illustrates the original defect. As one can see, the reconstructed defects cannot approximate the sides of the defect well, suffering from a lack of information due to missing signal peaks in the clusters. Moreover, with increasing noise, the approximation is stretched in the y -direction. However, with respect to the noise level, all reconstructions contain suitable information about the position and size of the defect.

Now we analyse the overall runtime of our reconstruction method. Table 9.13 shows the runtime in seconds for all three steps of the inversion method as well as the overall runtime. The calculations have been performed in Matlab 7 on a Macbook Pro (Mid

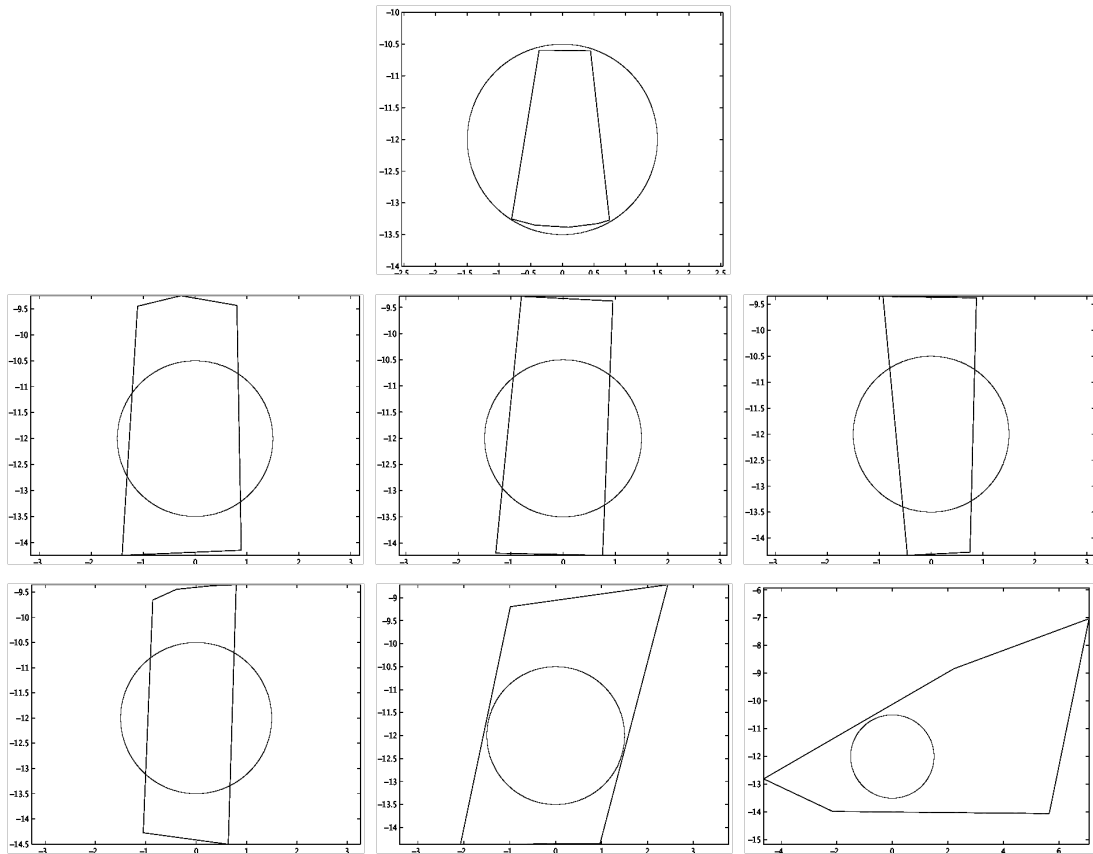


Fig. 9.23: Reconstruction of the defect in the weld plane with different noise levels sorted by increasing order of noise.

2009), 2.66 GHz Intel Core 2 Duo, 8 GB 1067 MHz DDR3 running OS X 10.8.4. Table 9.13 illustrates nicely that the reconstruction method introduced here can be used for real time reconstruction since the overall runtime is far below the time needed for the measurements. Note that while the deconvolution and clustering process have to be applied for each a-scan, the final reconstruction step only has to be computed for each cluster found. Moreover, one can reconstruct several defects without increase of the runtime by using a parallel implementation.

Noise	Deconvolution	Clustering	Reconstruction	overall runtime
0	0.1692	0.0820	1.2313	1.4825
0.005	0.2676	0.0782	1.6666	2.0124
0.01	0.2834	0.0744	1.2419	1.5997
0.015	0.2744	0.0744	1.4454	1.7942
0.02	0.3127	0.0818	1.2639	1.6584
0.025	0.1462	0.0124	1.7435	1.9020
0.035	0.3310	0.0297	2.3862	2.7469

Table 9.13: Runtime in seconds of all three inversion steps for different noise levels.

At the end of this example, we compare the reconstruction results with the approximations obtained with the SAFT algorithm described in Subsection 3.1.2. Therefore, we discretize the weld plane into a Cartesian grid and calculate the sum of corresponding amplitudes for each point. The runtime for SAFT with different grid step sizes is shown in Table 9.14, note that it is independent from the noise level since each amplitude is considered anyway. Figure 9.24 illustrates the reconstructions for the different noise levels. First we note that the reconstruction quality of SAFT is mostly independent from the noise level. This is due to its heuristic idea to suppress the noise by destructive interference. Hence, the SAFT approximation is slightly better for high noise. However, SAFT is no real reconstruction method but can only compute an image of the weld plane. This image also contains the components surface, its back wall and other signals caused by noise. Thus, a post-processing method is required that can calculate the defect form out of the given weld image. This may be a crucial step in the reconstruction process and hence more influenced by the noise level. Moreover, since SAFT maps back the ultrasonic signals to the weld plane, its vertical resolution is limited by the wave length. A comparison of the runtime of both algorithms shows that our inversion method is faster in most cases. Only for a low resolution SAFT might perform faster if the data is noisy and contains at least one defect cluster.

Step size	0.1	0.075	0.05	0.025
Runtime	0.9004	2.1019	5.7298	23.2237

Table 9.14: Runtime of SAFT for different grid step sizes.

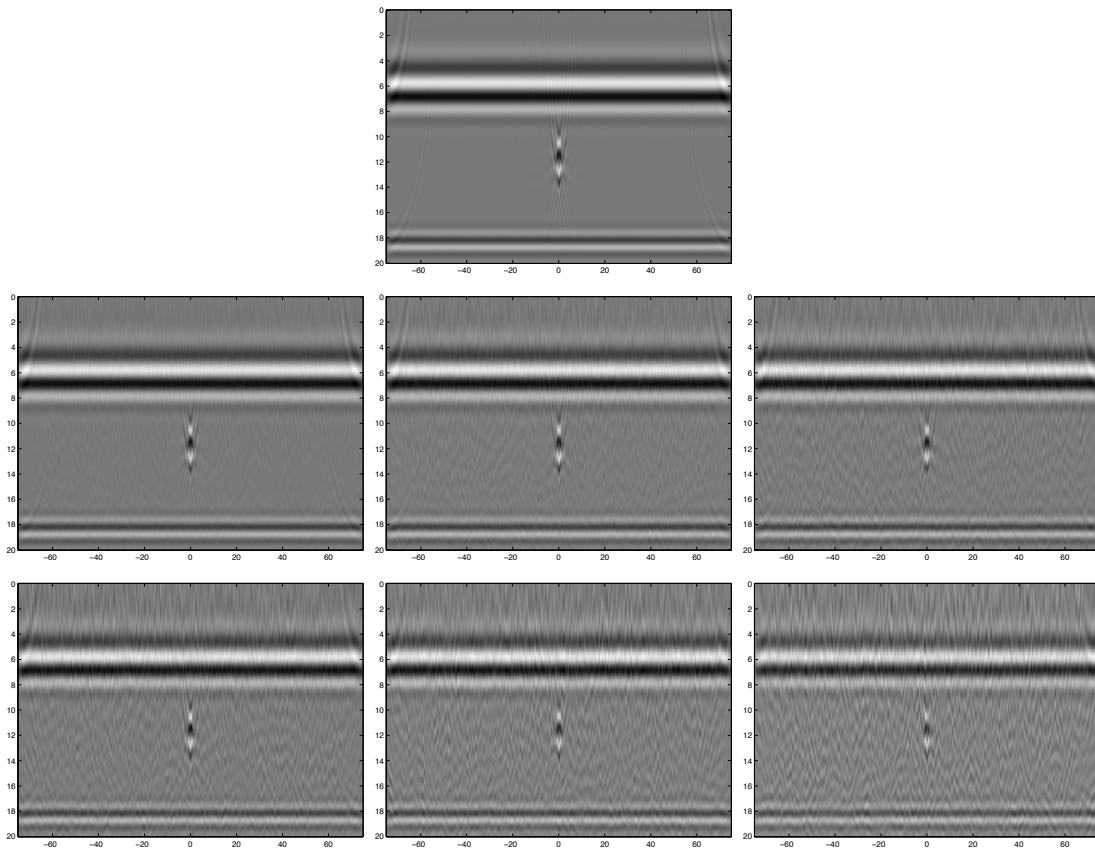


Fig. 9.24: SAFT reconstruction of the defect in the weld plane with different noise levels sorted by increasing order of noise.

9.4.2 Real ToFD Data

In our next example we apply our reconstruction method to real ToFD data shown in Figure 9.25. The data has been obtained from a sample of a large-diameter pipe (outer diameter 1066mm, wall thickness 23.3mm). For the left data, the weld seam has been tested with a ToFD system (Olympus Omniscan iX) with a 5MHz transducer, 6mm diameter (Olympus C543-SM) while for the right dataset a 10MHz transducer, 6mm diameter (Olympus C563-SM) with the same system has been used. Both transducers were applied with a wedge with 70° angle of incidence. The data contains two pore defects (left) and a lack of fusion at the end of the pipe, where the last part of the weld seam has been ground (right). All a-scans have an unknown non-linear sampling grid but we assume a resolution of 0.01 microseconds in time. Moreover, an a-scan was taken every 0.5mm.

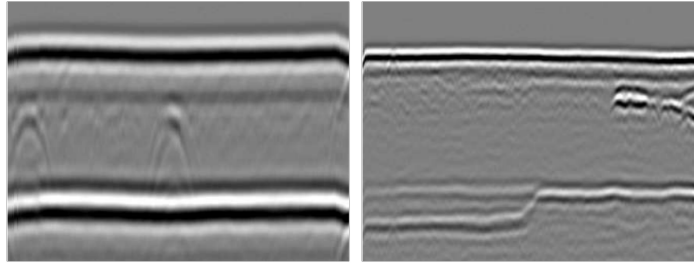


Fig. 9.25: ToFD real data including two pore defects (left) and a lack of fusion (right).

We use the Newton-HOMP method with the same parameters as in the example before to reconstruct the impulse function. The obtained results are shown in Table 9.15. Afterwards we apply the HOMP method to deconvolve the data. Thereby we use the discretization distance 0.01 and the parameters $L = 5$, $p = 5$, $\varepsilon_0 = 0.3$, $\varepsilon_1 = 0.05$ and $\varepsilon_2 = 0.15$. For the data containing the pore defects we use no weights while for the lack of fusion we apply the weights described in (9.1). The deconvolved data is shown in Figure 9.26. Next, we use the DBSCAN algorithm with $P_{\min} = 2$, $\gamma = 5.92^2$ and $\varepsilon = 1.54$ to estimate the defect clusters shown in Figure 9.27 where the first two clusters correspond to the pore defects and the last cluster represents the lack of fusion.

Dataset	ρ	ψ	ϕ
Pores	10.0653	2.3397	1.6061
Lack of fusion	25.5838	4.8050	0.5951

Table 9.15: Impulse parameter estimation results for real data shown in Figure 9.25.

Now we apply our last reconstruction step to each cluster. First, we recover a center point C using the heuristic approach 8.1.3 with $\varepsilon = 0.5$ and $\delta = 0$. Table 9.16 presents the reconstructed points where we assume the point $(0, 0, 0)$ to be in the upper left corner of the weld plane. Given these approximations of the center point we use the

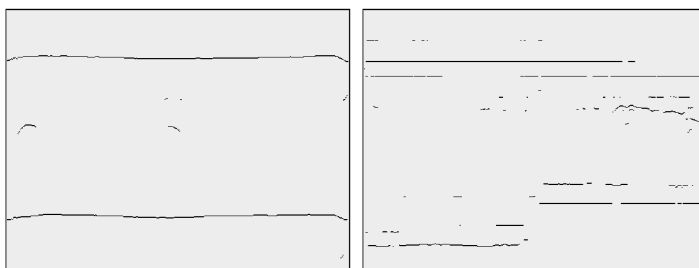


Fig. 9.26: Non-zero elements of the deconvolved data.

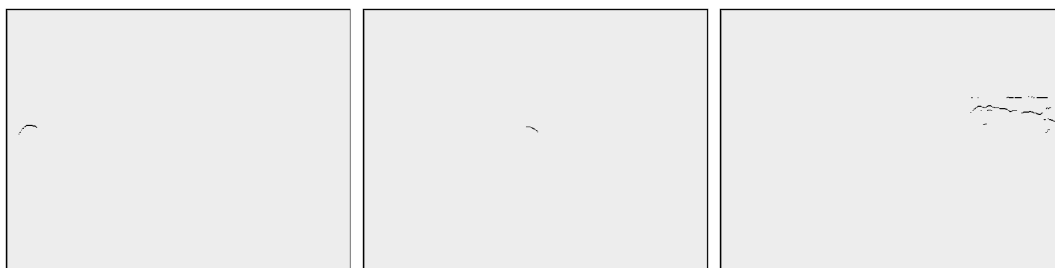


Fig. 9.27: Clusters obtained with DBSCAN: 2 clusters representing the pore defects in the first data set (left/middle) and one cluster corresponding to the lack of fusion in the second dataset.

spiderweb algorithm to reconstruct the defects where we choose an ε -separation with $\varepsilon = 0.001$. The reconstructed defects are illustrated in Figure 9.28 and the total runtime is listed in Table 9.17. Note that we do not have any information about the real defects inside the tube and thus cannot make a statement about the approximation properties. However, with the results obtained in the previous examples we suppose that the defect is reconstructed at an approximately right position. Moreover, the pore reconstructions may be smaller than the original defect since their clusters contain only view points while the lack of fusion is certainly be stretched in the vertical direction due to the noise.

Cluster	Pore (left side)	Pore (middle)	Lack of fusion
C	(13.2500,15.7763)	(107.75,15.9413)	(146.6290,13.1111)

Table 9.16: Reconstructed center points C for all three clusters.

Cluster	Deconvolution	Clustering	Reconstruction	overall runtime
Pore (left side)	0.2118	0.0727	1.9731	2.2576
Pore (middle)	0.2118	0.0727	0.9849	1.2694
Lack of fusion	0.3411	0.0713	5.3106	5.723

Table 9.17: Runtime in seconds of all three inversion steps for different noise levels.

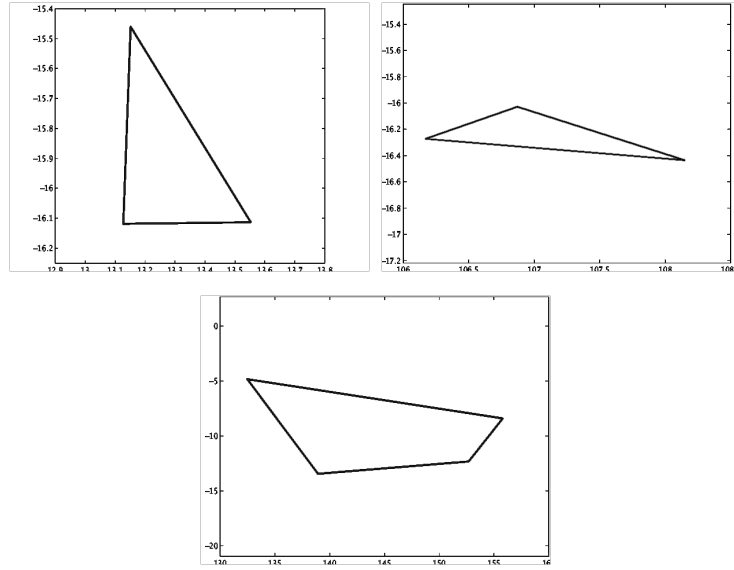


Fig. 9.28: Reconstructed defects: two pores (top) and a lack of fusion (bottom); units in mm.

9.4.3 Simulated Back Wall Defect

In our next example we consider a wall thickness measurement set-up with a simulated elliptic defect that lies diagonal in the back wall. More precisely, we assume a 20mm thick steel component with its surface in the $y = 0\text{mm}$ plane and the back wall in the $y = -20\text{mm}$ plane. The back wall has a defect in it that has its highest value $y = -17\text{mm}$ along a line between the points $(x, z) = (20\text{mm}, 20\text{mm})$ and $(x, z) = (50\text{mm}, 75\text{mm})$ and then decreases rapidly until it reaches $y = -20\text{mm}$. We use an ε -discretization with $\varepsilon = 0.5\text{mm}$, the defect is illustrated in Figure 9.29. We simulate the obtained ultrasound data for one probe being emitter and receiver at once with a focussing direction $(0, -1, 0)$ aiming directly onto the back wall. The angle of aperture is set to 12 degrees and the emitted impulse has a bandwidth of 40Mhz^2 , a frequency of 7Mhz and a shift of $\pi/2$. The attenuation coefficient is set to $\gamma = 0.0196 = (\psi/50)^2$. The a-scans were measured on a grid with grid size 0.5mm starting at $(x, y) = (5\text{mm}, 5\text{mm})$ and stopping at $(x, y) = (65\text{mm}, 90\text{mm})$. The a-scan itself has a sampling discretization of $0.01\mu\text{s}$. We used Gaussian noise with different variances $0, 0.001, 0.01, 0.02, 0.035, 0.05$ multiplied with the strongest back wall amplitude. The simulated data is three-dimensional as the probe moves along the surface in x - and z -direction. Figure 9.30 shows the obtained b-scan for $x = 35\text{mm}$ and different noise levels where we have cut off the surface reflection for a better contrast.

We use the Newton-HOMP method on the first 25 a-scans to calculate the impulse parameters where we choose the weights $w_k = 1$, $p = 2$, $\varepsilon_0 = \infty$, $\varepsilon_1 = \varepsilon_2 = 0$. Furthermore, we cut off the surface reflection and set $L = 1$, i.e. we seek for a 1-sparse representation. The obtained parameters are listed in Table 9.18. Afterwards we use the HOMP algo-

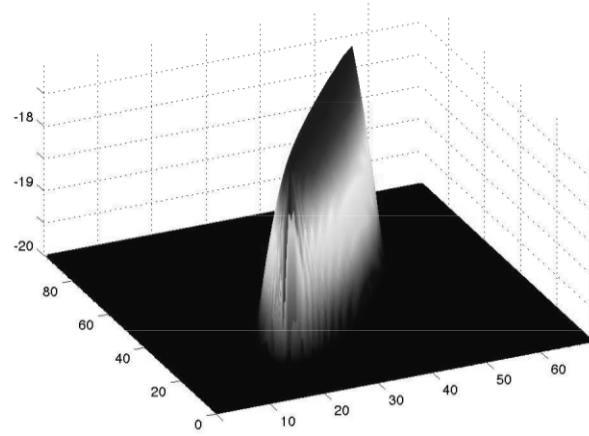


Fig. 9.29: Simulated elliptical back wall defect.

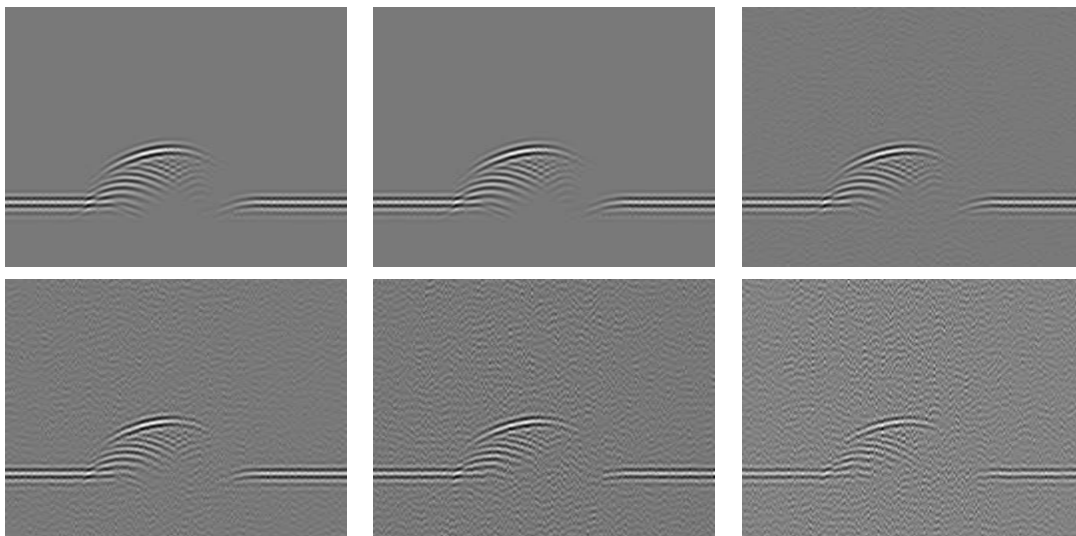


Fig. 9.30: Simulated b-scan for $x = 35\text{mm}$ and variances 0, 0.001, 0.01, 0.02, 0.035, 0.05 multiplied with the strongest back wall amplitude.

rithm with a discretization $t_{k+1} - t_k = 0.01\mu s$, $L = 2$, $p = 5$, $\varepsilon_1 = 0.05$, $\varepsilon_2 = 0.15$ and decreasing $\varepsilon_0 = 1, 0.5, 0.3, 0.25, 0.2, 0.15$ to counteract the increasing noise. Furthermore, we use the weights

$$w_k = e^{(\psi/50)^2 ct_k}$$

that make up for attenuation since we know that the important information about the back wall is late in time and thus more influenced by attenuation. Figure 9.31 illustrates the deconvolved data for the b-scans with $x = 35\text{mm}$ shown in Figure 9.30.

Noise	ρ	ψ	ϕ
0	33.4728	6.6761	0.9132
0.001	33.4489	6.6795	0.9129
0.01	34.8525	6.5854	0.7261
0.02	32.1995	6.5482	0.8285
0.035	34.3915	6.3046	0.7315
0.05	24.0681	6.5576	0.8829

Table 9.18: Impulse parameter estimation results for the simulated data with different noise levels.

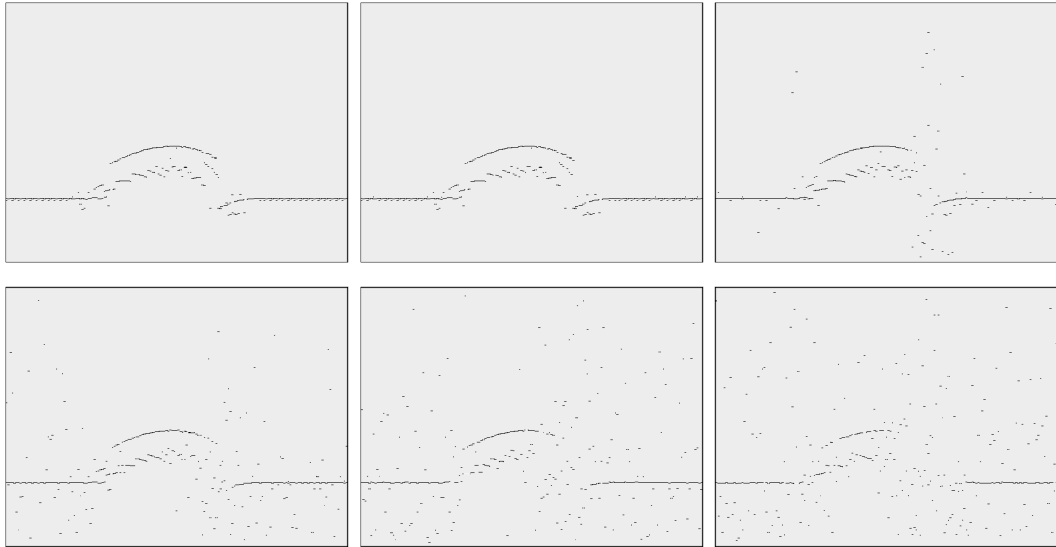


Fig. 9.31: Deconvolved b-scans for $x = 35\text{mm}$ and different noise levels.

Before we use our clustering algorithm to detect the defect signals, we reduce the amount of data by neglecting all found peaks that belong to the back wall. This can easily be done by counting the number of peaks found per time sampling point t_k . Figure 9.32 shows the number of points found for each index k for data with no noise (left) and with a noise level of 0.05 (right). As one can see, the index k' that corresponds to the back wall has by far the largest value. After this has been done, we use the adapted

DBSCAN method as before to calculate the defect cluster. Thereby we set $P_{\min} = 2$, $\gamma = 5.92^2$, $\varepsilon = 1.54$ for the four lower noise levels and $\varepsilon = 1.42$ for the datasets with the two highest noise levels. The obtained defect clusters are illustrated in Figure 9.33 where again some minor clusters were neglected.

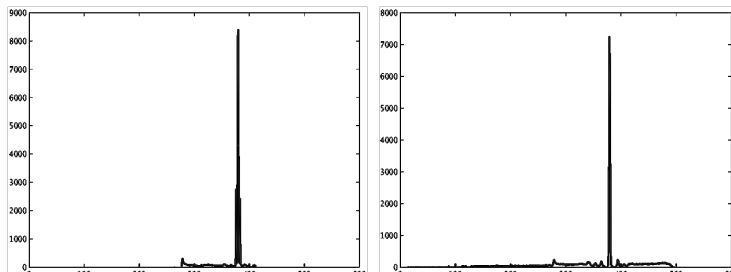


Fig. 9.32: Number of found peaks per index k corresponding to a time sampling point t_k .

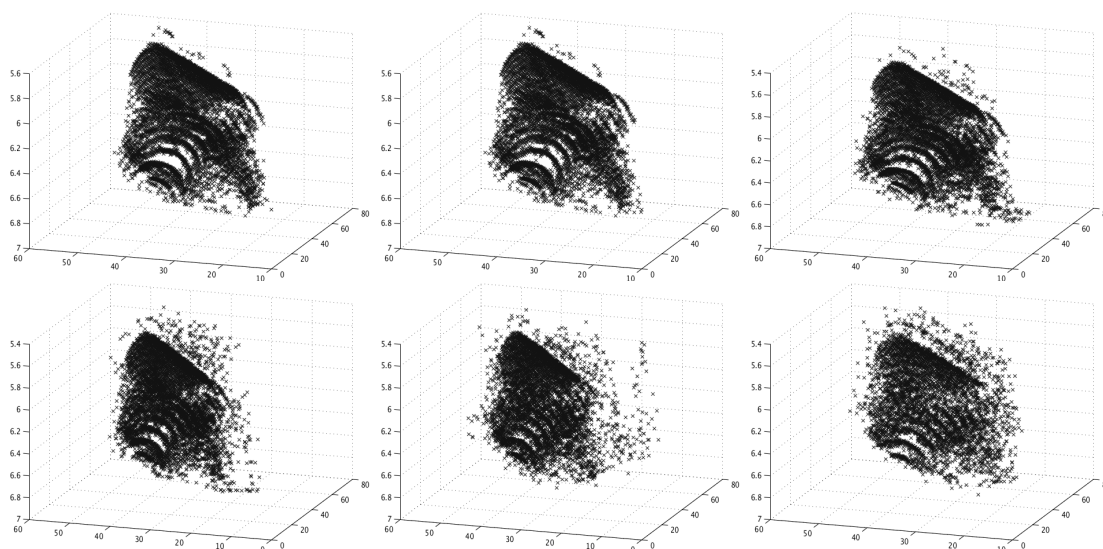


Fig. 9.33: Obtained clusters using DBSCAN on the deconvolved data for different noise levels where the back wall peaks are neglected; x - and y -axis in mm, z -axis in μs .

In this 3D-case the third reconstruction step is computationally challenging. Although our Algorithm 8.2.22 is designed to reconstruct both two- and three-dimensional defects, the number of sets M in $\{\hat{\zeta}_j\}_{j=1}^M = A \cap B$, where A is an ε -separation and B is adapted to the measurements, is drastically increasing for a suitable small ε . While the number of sets in A can be reduced by choosing ε larger, the crucial factor is the number of sets in B . It depends on ε as well as on the number of points in the cluster. While in the previous examples, a defect cluster usually consists of about 50 – 200 points, the here reconstructed clusters contain about 6000 peaks. In order to keep a reasonable computation time, we have to choose $\varepsilon \gg 0$ yielding only a rough approximation. However, we will overcome this problem by applying Algorithm 8.2.22 for each b-scan with fixed

position x of the data. The two-dimensional reconstructions are combined afterwards to obtain a three-dimensional approximation. Indeed, in this way the connection between neighbouring b-scans is lost but the algorithm can be performed in reasonable time and the approximation results are quite accurate. Using the heuristic approach 8.1.4 to approximate the center point C and choosing an ε -separation with $\varepsilon = 0.001$ the algorithm yields the reconstructions shown in Figure 9.34 as a point cloud and 9.35 as triangulation. As one can see, the top of the defect is reconstructed quite well depending on the noise level while the sides suffer from a loss of information. This is due to the fact that the sides are nearly parallel to the focussing direction. The total runtime of our inversion algorithm is listed in Table 9.19. Note that due to the huge number of points in a cluster, the runtime is quite long. However, as the data consist out of 121 b-scans, the runtime per b-scan is less then 2 seconds.

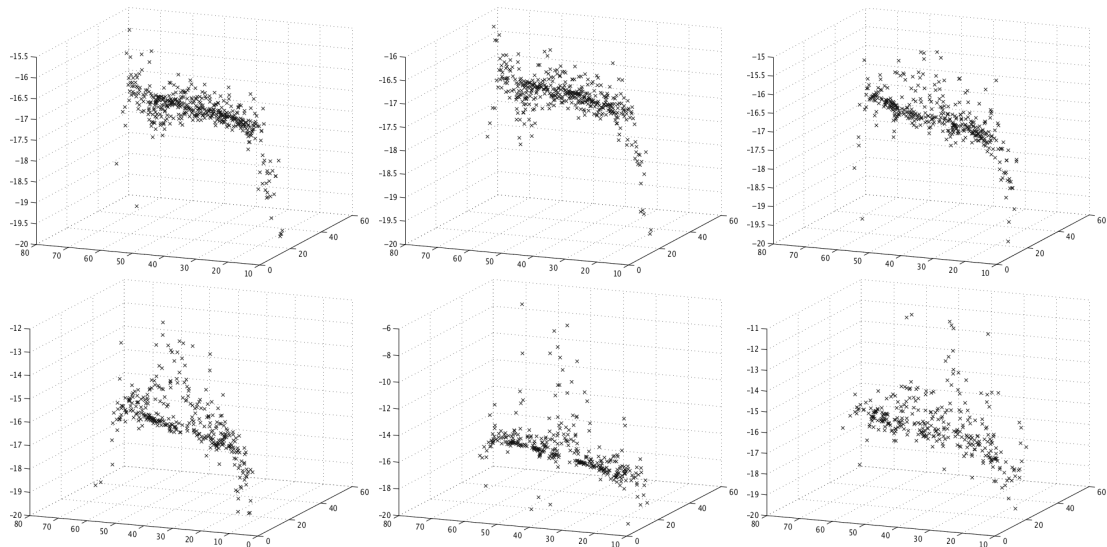


Fig. 9.34: Reconstructed defect for different noise levels as point cloud.

Noise	Deconvolution	Clustering	Reconstruction	overall runtime
0	51.6399	5.3277	182.0422	239.0098
0.001	33.6790	5.3718	182.2971	221.3479
0.01	38.9569	6.5460	188.2486	233.7515
0.02	35.0077	12.3092	191.7488	239.0657
0.035	39.3868	10.6761	173.0896	223.1525
0.05	20.3176	9.7962	148.7618	178.8756

Table 9.19: Runtime in seconds of all three inversion steps for different noise levels.

For comparison, we also performed SAFT on the original three-dimensional data. As our algorithm, SAFT also suffers from the huge amount of data. Hence, we set the discretization grid size to 0.5mm. Figure 9.36 shows the largest amplitudes of the reconstructed

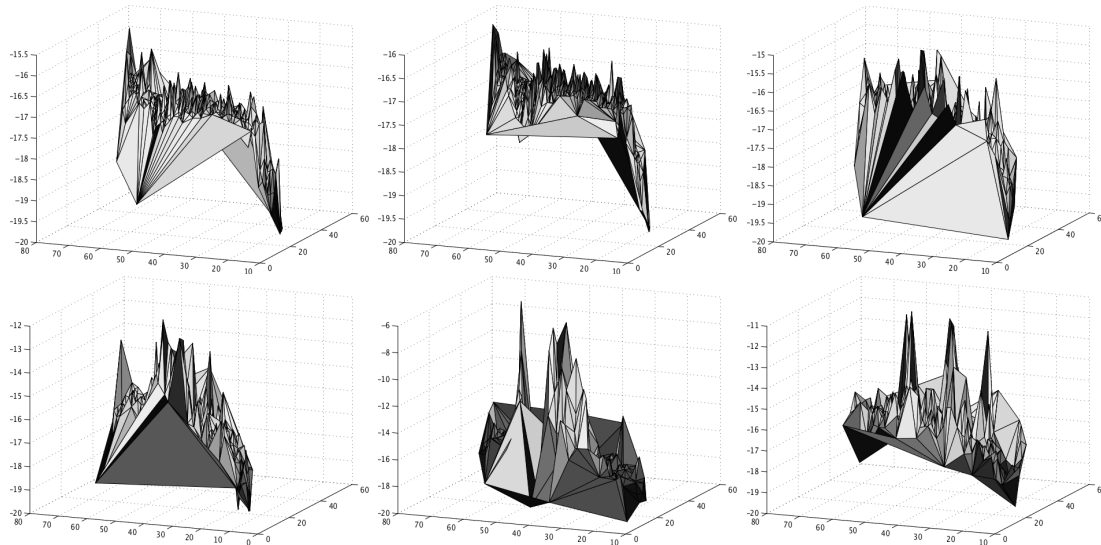


Fig. 9.35: Reconstructed defect for different noise levels as triangulation.

SAFT image. The calculation for each dataset took about 330 seconds. One can identify four structures in Figure 9.36. The largest structure at the top is the reconstruction of the surface. The structure at the bottom illustrates the back wall while the structure second from the bottom is the upper part of the defect. However, the structure second from the top seems to be an echo of the defect caused by the wave structure. Although SAFT has been applied to the three-dimensional data at once, the algorithm was not able to produce results comparable to our algorithm that only used the two-dimensional information.

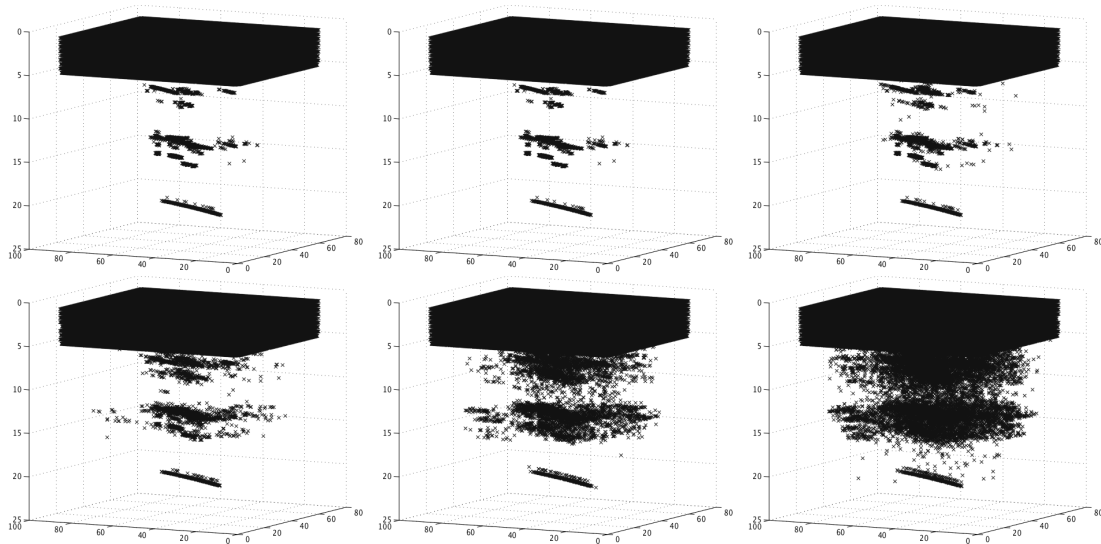


Fig. 9.36: Largest reconstructed amplitudes using SAFT on the 3d data.

9.4.4 Real Wall Thickness Measurement Data

As a last example in this chapter, we want to consider real wall thickness measurement data obtained from a 14mm thick steel tube. Figure 9.37 shows exemplary a b-scan of the data. As one can see, there is a defect in the back wall of the tube. An imprint of this defect has been produced, sized and digitalized, see Figure 9.38. Hence, we do know the form and size of the defect approximately.



Fig. 9.37: B-scan out of the given real data.

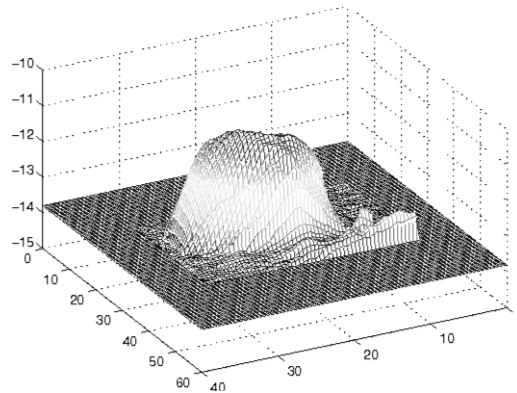


Fig. 9.38: Digitalized imprint of the back wall defect extended by zero outside.

Using the Newton-HOMP approach to reconstruct the impulse parameters, we obtain a bandwidth of 57.8695Mhz^2 , a frequency of 6.2916Mhz and a shift of 0.0228 . For deconvolution of the data, we use the HOMP algorithm with a discretization $t_{k+1} - t_k = 0.01\mu\text{s}$, $L = 2$, $p = 5$, $\varepsilon_0 = 0.75$, $\varepsilon_1 = 0.05$ and $\varepsilon_2 = 0.15$. Furthermore, we apply the weights $w_k = e^{(\psi/50)^2 ct_k}$ as in the last example. Figure 9.39 exemplarily illustrates the deconvolved data for the b-scan shown in Figure 9.37 where the surface reflection has been truncated. Again, before applying a clustering algorithm, we remove all back wall

peaks from the data. Then we reconstruct the defect cluster shown in Figure 9.40 using DBSCAN with $P_{\min} = 2$, $\gamma = 5.92^2$, $\varepsilon = 1.54$.

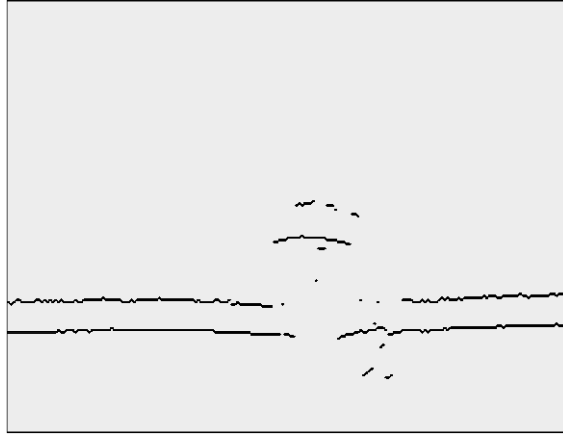


Fig. 9.39: Non-zero elements of the deconvolved data for the b-scan shown in Figure 9.37; the surface reflection has been truncated.

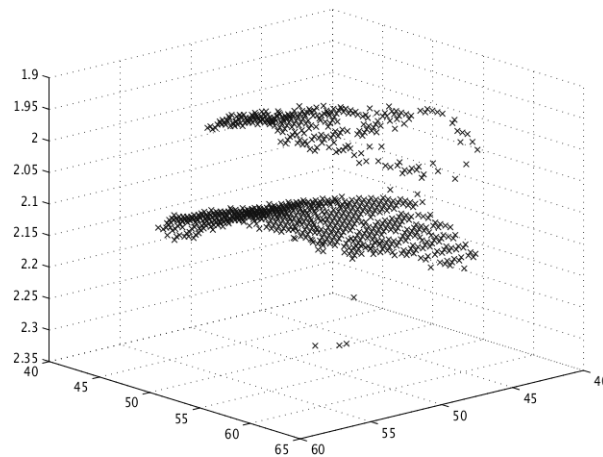


Fig. 9.40: Cluster corresponding to the back wall defect obtained with DBSCAN.

In a last step we reconstruct the defect applying Algorithm 8.2.22 on each b-scan. We use the heuristic method 8.1.4 and set $\varepsilon = 0.001$ for the ε -separation. The result is illustrated in Figure 9.41. Note that the sampling grid differs from that in Figure 9.38 since the imprint could be sampled more detailed. Moreover, although we may assume that Figure 9.38 shows an reasonable image of the back wall defect, it is not quite clear how big the approximation error due to the imprinting, sampling and digitalization is. Since the defect cluster in this example only contains about 600 points, the complete reconstruction of the defect took 29.8707 seconds, including all three steps of our inversion method.

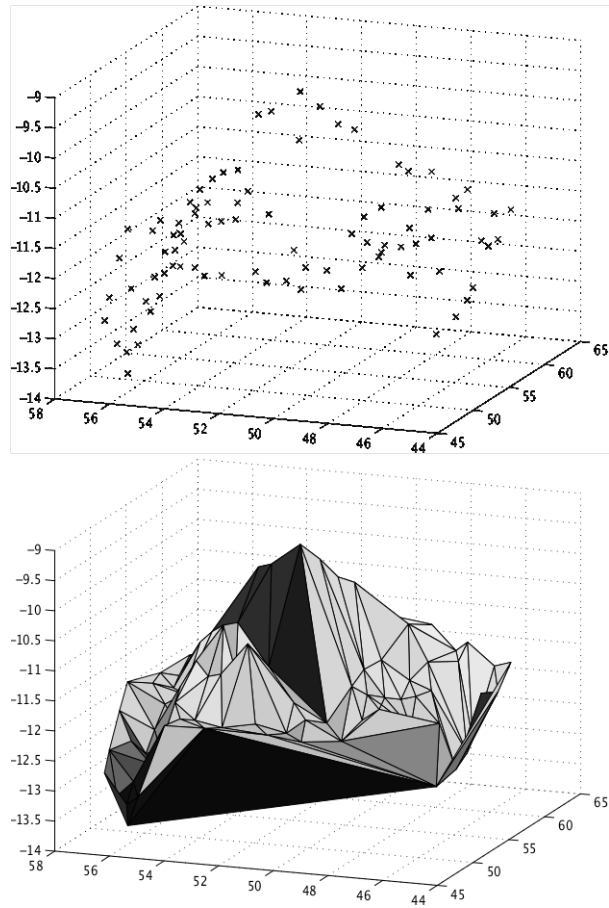


Fig. 9.41: Reconstructed defect as points cloud (top) and triangulation (bottom).

10 Conclusions

Let us summarize the results of this thesis. As we have learned from the first chapters, ultrasonic non-destructive testing is a complicated task with many unsolved problems. Surely, today the behaviour of ultrasound in material is well understood and many simulation and reconstruction methods are known. However, the industrial framework implicates the challenging issue to balance the resources of time and accurate information reconstruction. Due to the very fast production process both of these resources are strictly limited. In many cases, only few ultrasound measurements can be taken to analyse the component. This leads to inverse problems with highly incomplete data and to the task of finding a reconstruction method that is fast and stable.

In this thesis, we presented a new model based reconstruction approach. The introduced model was kept as simple as possible to allow a fast inversion algorithm. Nevertheless, a close connection to inverse scattering via the Born approximation was shown. Based on this model, we were able to introduce our reconstruction idea that mainly consists of three parts: Extraction of information, assignment of information and reconstruction. The great benefit of this idea is to use the valuable resource time only as much as necessary. In a first step, the amount of input data is drastically reduced and only important information is kept. This allows the following steps to operate on a small amount of data admitting a vast decrease of runtime. The second step classifies the data into different groups such as noise, back wall, lateral wave or defect. Such a classification is crucial for real time computation. The third step is the actual reconstruction of the defect. The method used for this step is often time consuming as the defect reconstruction is a complicated problem. Hence it is important that the amount of data is reduced by the first two steps of our reconstruction idea. Thus, the time expensive operations are only computed if needed. Independent of the here presented explicit method for those three steps, the basic idea behind this method is a fundamental step to overcome the problem of high dimensional data. Today's known methods such as SAFT are suffering from this high dimensionality with a high computational effort and thus cannot be performed in real time especially in the 3D case.

After having introduced the basic idea of our method, we presented algorithms for all three reconstruction steps. We used a sparse deconvolution method to extract the important information out of the input data. Thereby, we made special use of so called Greedy methods which often have a very low complexity. As it was shown in the numerical examples, the introduced HOMP algorithm is even able to handle data with high noise levels. Furthermore, an extension of this approach, the Newton-HOMP algorithm, is also able to reconstruct the unknown impulse parameters. For the second

step, we used clustering algorithms to classify the data. More precisely, an adaption of DBSCAN was used, an algorithm that was especially designed to handle noisy data. In addition, we also introduced DBCLAN, another adaption of DBSCAN that is able to separate overlapping clusters which may appear if two defects are close to each other. The numerical results show that both algorithms are capable of classifying the data into different clusters, here the DBCLAN was more prone to noise than DBSCAN. For our last step we introduced a completely new reconstruction method based on a simple algorithm that we call spiderweb algorithm. This method is able to reconstruct an object, or part of it, from given distances between the object's boundary and positions outside the object. Thereby, it automatically adapts to the given information and reconstructs those parts of the object where information is given while the other part is neglected. Another advantage of this method is that it directly reconstructs a number of points representing the defect while e.g. SAFT can only reconstruct an image of the component that has to be interpreted. The stability of our new method has not only been confirmed by the numerical examples shown in this thesis but we were also able to derive several theoretical results that prove approximation properties for this algorithm.

Altogether, the presented methods combined to a reconstruction technique form a fast and stable algorithm. The numerical examples have shown that the method is able to beat SAFT in both runtime and approximation accuracy. Although the individual methods can surely be improved, the ideas of data reduction and defect reconstruction as point cloud are an important step forward to real time computation and defect reconstruction. Thus, the algorithm can be seen as construction kit for a new type of inversion methods where each of the three parts can be exchanged, improved or adapted according to the underlying application.

Bibliography

- [1] O. Johannsen. *Geschichte des Eisens*. Verl. Stahleisen Düsseldorf, 1953.
- [2] G. Krankenhagen and H. Laube. *Wege der Werkstoffprüfung*. Deutsches Museum München, 1979.
- [3] J. Krautkrämer and H. Krautkrämer. *Werkstoffprüfung mit Ultraschall*. Springer Berlin, 1986.
- [4] J. Krautkrämer and H. Krautkrämer. *Ultrasonic Testing of Materials*. Springer Berlin, 1990.
- [5] B. M. Lempriere. *Ultrasound and Elastic Waves - Frequently Asked Questions*. Academic Press Amsterdam, 2002.
- [6] W. Demtröder. *Experimentalphysik 1 - Mechanik und Wärme*. Springer Berlin, 2013.
- [7] K. J. Langenberg, R. Marklein, and K. Mayer. *Theoretische Grundlagen der zerstörungsfreien Materialprüfung mit Ultraschall*. Oldenbourg München, 2009.
- [8] U. Schlengermann. *Taschenbuch Ultraschall-Werkstoffprüfung*. Vulkan Essen, 1992.
- [9] A. Leger and M. Deschamps. *Ultrasonic Wave Propagation in Non Homogeneous Media*, volume 128 of *Springer Proceedings in Physics*. Springer Berlin, 2009.
- [10] F. Boßmann. Entwicklung einer automatisierten Auswertung von bildgebenden Ultraschallverfahren. Diploma thesis, University of Duisburg-Essen, 2009.
- [11] C. Huygens. *Treatise on Light*. University of Chicago Press, translated edition, 1912.
- [12] H. Ammari, editor. *Mathematical Modeling in Biomedical Imaging II*. Springer Berlin, 2012.
- [13] M. Vorländer. *Auralization*. Springer Berlin, 2008.
- [14] M. Bruneau. *Fundamentals of Acoustics*. ISTE London, translated by T. Scelo edition, 2006.
- [15] H. Kuttruff. *Physik und Technik des Ultraschalls*. S. Hirzel Stuttgart, 1988.

- [16] S. Mondal and T. Sattar. An overview ToFD method and its mathematical model. *NDT.net*, 5(4), 2000.
- [17] N. Trimborn. The Time-of-Flight-Diffraction-Technique. *NDT.net*, 2(09), 1997.
- [18] K. G. Prabhakaran, B. S. Wong, and Y. Y. Teng. Time of Flight Diffraction - an alternate non-destructive testing procedure to replace traditional methods. *NDT.net*, 10(5), 2005.
- [19] B. Hofmann. *Mathematik inverser Probleme*. Teubner Stuttgart Leipzig, 1999.
- [20] A. van den Biggelaar et al. Automated evaluation for ToFD. *ECNDT*, 2010.
- [21] M. Spies and H. Rieder. Quantitative Fehlercharakterisierung aus Ultraschall-HF-Daten mittels simulationsgestützter Auswerteverfahren. *Dach-Jahrestagung, St. Gallen*, 2008.
- [22] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Pearson New Jersey, 2008.
- [23] G. Vogt and H. Zück. Hightech auf Schienen: Die automatisierte Schienenprüfung mittels Ultraschall am Beispiel des Schienenprüfzuges der Deutsche Bahn Netz AG. *NDT.net*, 9(10), 2004.
- [24] Olympus. Ultrasonic flaw detectors. <http://www.olympus-ims.com/en/ut-flaw/>.
- [25] GE Measurement & Control. Portable flaw detectors. <http://www.ge-mcs.com/en/ultrasound/portable-flaw-detectors.html>.
- [26] M. Jobst and G. D. Connolly. Demonstration of the application of the Total Focussing Method to the inspection of steel welds. *ECNDT*, 2010.
- [27] H. Rieder, A. Dillhöfer, M. Spies, A. Graff, T. Orth, and T. Kersting. SAFT- und ToFD-Auswertung für die Ultraschall-Schweißnahtprüfung von längsnahtgeschweißten Großrohren. *DGZfP-Jahrestagung*, 2010.
- [28] M. Spies and W. Jager. Synthetic aperture focusing for defect reconstruction in anisotropic media. *Ultrasonics*, 41(2):125–131, 2003.
- [29] A. Velichko and P. D. Wilcox. An analytical comparison of ultrasonic array imaging algorithms. *J. Acoust. Soc. Am.*, 127(4):2377–2384, 2010.
- [30] R. Marklein, J. Miao, M. Rahman, and K. J. Langenberg. Inverse scattering and imaging in NDT: Recent applications and advances. *ECNDT*, 2006.
- [31] A. J. Devaney. Super-resolution processing of multi-scale data using time reversal and MUSIC. *Preprint at <http://www.ece.neu.edu/faculty/devaney/ajd/preprints.htm>*.

-
- [32] E. G. Asgedom, L.-J. Gelius, A. Austeng, S. Holm, and M. Tygel. Time-reversal multiple signal classification in case of noise: A phase-coherent approach. *J. Acoust. Soc. Am.*, 130(4):2024–2034, 2011.
- [33] F. K. Gruber, E. A. Marengo, and A. J. Devaney. Time-reversal imaging with multiple signal classification considering multiple scattering between the targets. *J. Acoust. Soc. Am.*, 115(6):3042–3047, 2004.
- [34] R. Rojas. *Neural Networks*. Springer Berlin, 1996.
- [35] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford Univ. Press New York, 2005.
- [36] R. Polikar, L. Udpa, and S. S. Udpa. Frequency invariant classification of ultrasonic weld inspection signals. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 45(3):614–625, 1998.
- [37] S. Lalithakumari, B. Sheelarani, and B. Venkatraman. Classification of ToFD signals by artificial neural network. *WCNDT*, 2012.
- [38] A. R. Filho, A. Petraglia, and M. R. Petragli. Automatic welding diagnosis using neural network. *NDT.net*, 5(4), 2000.
- [39] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*, volume 93 of *Applied Mathematical Sciences*. Springer New York, 1998.
- [40] F. Cakoni and D. Colton. *Qualitative Methods in Inverse Scattering Theory - An Introduction*. Springer Berlin, 2006.
- [41] K. Chadan, D. Colton, L. Päiväranta, and W. Rundell. *An Introduction to Inverse Scattering and Inverse Spectral Problems*. SIAM Philadelphia, 1997.
- [42] J. L. Davis. *Mathematics of Wave Propagation*. Princeton University Press New Jersey, 2000.
- [43] T. S. Angell, R. E. Kleinman, and F. Hettlich. The resistive and conductive problems for the exterior Helmholtz equation. *SIAM J. Appl. Math.*, 50(6):1607–1622, 1990.
- [44] K. J. Langenberg. *Basic Methods of Tomography and Inverse Problems*. IOP Bristol, 1987.
- [45] R. Kress. Uniqueness and numerical methods in inverse obstacle scattering. *J. Phys.: Conf. Ser.*, 73(1):012003, 2007.
- [46] R. Potthast. Sampling and probe methods - an algorithmical view. *Computing (Springer)*, 75(2-3):215–235, 2005.
- [47] D. Colton and R. Kress. Using fundamental solutions in inverse scattering. *Inverse Probl.*, 22(3):R49, 2006.

- [48] P. M. P. de Carvalho Serranho. *A Hybrid Method for Inverse Obstacle Scattering Problems*. Phd thesis, Georg-August-University Göttingen, 2007.
- [49] S. J. Norton. Iterative inverse scattering algorithms: Methods of computing frechet derivatives. *J. Acoust. Soc. Am.*, 106(5):2653–2660, 1999.
- [50] T. Hohage. *Iterative Methods in Inverse Obstacle Scattering: Regularization Theory of Linear and Nonlinear Exponentially Ill-Posed Problems*. Universitätsverlag Rudolf Traumer Linz, 1999.
- [51] F. Bauer, T. Hohage, and A. Munk. Iteratively regularized Gauss-Newton methods for nonlinear inverse problems with random noise. *SIAM J. Numer. Anal.*, 47(3):1827–1846, 2009.
- [52] H. Harbrecht and T. Hohage. Fast methods for three-dimensional inverse obstacle scattering problems. *J. Int. Eq. Appl.*, 19(3):237–260, 2007.
- [53] R. Potthast. A fast new method to solve inverse scattering problems. *Inverse Probl.*, 12(5):731–742, 1996.
- [54] R. Potthast. A point-source method for inverse acoustic and electromagnetic obstacle scattering problems. *IMA J. Appl. Math.*, 61(2):119–140, 1998.
- [55] D. Colton and A. Kirsch. A simple method for solving inverse scattering problems in the resonance region. *Inverse Probl.*, 12(4):383–394, 1996.
- [56] F. Schmidt, T. Hohage, R. Klose, A. Schädle, and L. Zschiedrich. Pole condition: A numerical method for Helmholtz-type scattering problems with inhomogeneous exterior domain. *J. Comput. Appl. Math.*, 218(1):61–69, 2008.
- [57] T. Hohage. On the numerical solution of a three-dimensional inverse medium scattering problem. *Inverse Probl.*, 17(6):1743–1763, 2001.
- [58] D. R. Luke and R. Potthast. Image processing for limited aperture inverse acoustic obstacle scattering. *Pacific Institute for the Mathematical Sciences (preprint)*, 2003.
- [59] D. R. Luke. Multifrequency inverse obstacle scattering: the point source method and generalized filtered backprojection. *Math Comput. Simul.*, 66(4):297–314, 2004.
- [60] G. Bao and P. Li. Inverse medium scattering problems in near-field optics. *J. Comput. Math.*, 25(3):252–265, 2007.
- [61] M. Ikehata, E. Niemi, and S. Siltanen. Inverse obstacle scattering with limited-aperture data. *Inverse Probl. Imaging*, 6(1):77–94, 2012.
- [62] M. Hanke. One shot inverse scattering via rational approximation. *SIAM J. Imaging Sci.*, 5(1):465–482, 2012.
- [63] J. Göbel. *Radartechnik*. VDE Verlag Berlin, 2011.

-
- [64] B.-C. Wang. *Digital Signal Processing Techniques and Applications in Radar Image Processing*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2008.
- [65] R. E. Hansen. Introduction to synthetic aperture sonar. *Intech online resource*, <http://www.intechopen.com/books/sonar-systems/introduction-to-synthetic-aperture-sonar>, 2011.
- [66] H. Ammari, editor. *Mathematical Modeling in Biomedical Imaging I*. Springer New York, 2009.
- [67] L. L. Bonilla, editor. *Inverse Problems and Imaging*. Springer Berlin, 2008.
- [68] W. Demtröder. *Experimentalphysik 2 - Elektrizität und Optik*, volume 4. Springer Berlin, 2006.
- [69] S. G. Lipson, H. Lipson, and D. S. Tannhauser. *Optical Physics*. Cambridge University Press Cambridge, 1995.
- [70] R. Dittion. *Modern Geometrical Optics*. John Wiley & Sons, Inc., New York, 1998.
- [71] U. Harten. *Physik - Eine Einführung für Ingenieure und Naturwissenschaftler*. Springer New York, 2011.
- [72] D. T. Blackstock. *Fundamentals of Physical Acoustics*. John Wiley & Sons, Inc., New York, 2000.
- [73] A. Schoch. *Schallreflexion, Schallbrechung und Schallbeugung*, volume 23 of *Ergebnisse der Exakten Naturwissenschaften*. Springer Berlin, 1950.
- [74] R. Demirli and J. Saniie. Model-based estimation of ultrasonic echoes part I: Analysis and algorithms. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 48(3):787–802, 2001.
- [75] F. Natterer. An error bound for the Born approximation. *Inverse Probl.*, 20(2):447–452, 2004.
- [76] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.*, 41(12):3397–3415, 1993.
- [77] J. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Trans. Inf. Theory*, 50(10):2231–2242, 2004.
- [78] G. Prony. Essai expérimental et analytique sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l’eau et de la vapeur de l’alkool, à différentes températures. *J. de l’École Polytechnique*, 1:24–76, 1795.
- [79] T. Peter, D. Potts, and M. Tasche. Nonlinear approximation by sums of exponentials and translates. *SIAM J. Sci. Comput.*, 33(4):1920–1944, 2011.

- [80] F. Boßmann, G. Plonka, T. Peter, O. Nemitz, and T. Schmitte. Sparse deconvolution methods for ultrasonic NDT. *Journal of Nondestructive Evaluation*, 31(3):225–244, 2012.
- [81] A. C. Fannjiang. The MUSIC algorithm for sparse objects: A compressed sensing analysis. *Inverse Probl.*, 27(3):035013, 2011.
- [82] D. L. Donoho. For most large underdetermined systems of linear equations the minimal l_1 -norm solution is also the sparsest solution. *Commun. Pure Appl. Math.*, 59(6):797–829, 2006.
- [83] T. H. Cormen, C. E. Leiserson, R. Rivest, and C. Stein. *Algorithmen - Eine Einführung*, volume 2. Oldenbourg München, 2007.
- [84] V. Temlyakov. *Greedy Approximation*. Cambridge University Press New York, 2011.
- [85] N. Ruiz-Reyes, P. Vera-Candeas, J. Curpián-Alonso, R. Mata-Campos, and J.C. Cuevas-Martinez. New matching pursuit-based algorithm for SNR improvement in ultrasonic NDT. *NDT & E International*, 38(6):453–458, 2005.
- [86] N. Ruiz-Reyes, P. Vera-Candeas, J. Curpián-Alonso, J.C. Cuevas-Martinez, and J.L. Blanco-Claraco. High-resolution pursuit for detecting flaw echoes close to the material surface in ultrasonic NDT. *NDT & E International*, 39(6):487–492, 2006.
- [87] E. Mor, A. Azoulay, and M. Aladjem. A matching pursuit method for approximating overlapping ultrasonic echoes. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 57(7):1996–2004, 2010.
- [88] D. Christensen, S. Das, and A. N. Srivastava. Highly scalable matching pursuit signal decomposition algorithm. *Proc. IWSHM*, pages 1194–1201, 2009.
- [89] B. Mailhé, R. Gribonval, F. Bimbot, and P. Vanderghenst. A low complexity orthogonal matching pursuit for sparse signal approximation with shift-invariant dictionaries. *ICASSP*, pages 3445–3448, 2009.
- [90] E. Liu and V. N. Temlyakov. The orthogonal super greedy algorithm and applications in compressed sensing. *IEEE Trans. Inf. Theory*, 58(4):2040–2047, 2012.
- [91] K. F. Kaaresen and E. Bolviken. Blind deconvolution of ultrasonic traces accounting for pulse variance. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 46(3):564–573, 1999.
- [92] J. M. Mendel. *Optimal seismic deconvolution: An estimation-based approach*. Academic Press New York, 1983.
- [93] A. T. Walden. Non-Gaussian reflectivity, entropy and deconvolution. *Geophysics*, 50(12):2862–2888, 1985.

- [94] H.-H. Ciang and C. L. Nikias. Adaptive deconvolution and identification of non-minimum phase FIR systems based on cumulants. *IEEE Trans. Automatic Contr.*, 35(1):36–47, 1990.
- [95] A. K. Nandi, D. Mämpel, and B. Roscher. Comparative study of deconvolution algorithms with applications in non-destructive testing. *IEEE Dig.*, 145(1):1–6, 1995.
- [96] R. Neelamani, H. Choi, and R. Baraniuk. ForWaRD: Fourier-wavelet regularized deconvolution for ill-conditioned systems. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 52(2):418–432, 2004.
- [97] R. H. Herrera, R. Orozon, and M. Rodriguez. Wavelet-based deconvolution of ultrasonic signals in nondestructive evaluation. *J. Zhejiang Univ. SCIENCE A*, 7(10):1748–1756, 2006.
- [98] T. Olofsson. Computationally efficient sparse deconvolution of b-scan images. *Proc. IEEE Ultrason. Symp.*, 1:540–543, 2005.
- [99] T. Olofsson and E. Wennerström. Sparse deconvolution of b-scan images. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 54(8):1634–1641, 2007.
- [100] R. Demirli and J. Saniie. Model-based estimation of ultrasonic echoes part II: Nondestructive evaluation applications. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 48(3):803–811, 2001.
- [101] I. Ziskind and M. Wax. Maximum likelihood localization of multiple sources by alternating projection. *IEEE Trans. Acoust. Speech Signal Process.*, 36(10):1553–1560, 1988.
- [102] C. F. J. Wu. On the convergence properties of the EM algorithm. *Ann. Stat.*, 11(1):95–103, 1983.
- [103] P. J. Chung and J. F. Böhme. Comparative convergence analysis of EM and SAGE algorithms in DOE estimation. *IEEE Trans. Signal Process.*, 49(12):2940–2949, 2001.
- [104] J. Bacher, A. Pöge, and K. Wenzig. *Clusteranalyse - Anwendungsorientierte Einführung in Klassifikationsverfahren*. Oldenbourg München, 2010.
- [105] N. A. Rathod and S. A. Ladhake. Review of data clustering algorithms. *IJCIS*, 2(3):56–58, 2011.
- [106] R. C. T. Lee. Clustering analysis and its applications. *Advances in Information Systems Science*, 6:169–292, 1981.
- [107] C. Elkan. Using the triangle inequality to accelerate k-means. *Proc. ICML*, pages 147–153, 2003.

- [108] K. Hwang, S. Mandayam, S. S. Udpa, L. Udpa, W. Lord, and M. Atzal. Characterization of gas pipeline inspection signals using wavelet basis function neural networks. *NDT & E International*, 33(8):531–545, 2000.
- [109] M. Ester, H. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD*, 96:226–231, 1996.
- [110] M. Kryszkiewicz and P. Lasek. TI-DBSCAN: Clustering with DBSCAN by means of the triangle inequality. *Lect. Notes Comput. Sci.*, 6086:60–69, 2010.
- [111] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press New York, 1981.
- [112] L. A. Zadeh. Fuzzy sets. *Inf. Control*, 8(3):338–353, 1965.
- [113] Z. Xiong, R. Chen, Y. Zhang, and X. Zhang. Multi-density DBSCAN algorithm based on density levels partitioning. *J. Inf. Comput. Sci.*, 9(10):2739–2749, 2012.
- [114] W. F. Demjanov and W. N. Malozemov. *Einführung in Minimax-Probleme*. Geest u. Portig Leipzig, 1973.
- [115] Jochen Fraatz. Numerische Behandlung von Minimax-Aufgaben mit der Methode der Lagrange-Multiplikatoren. Diploma thesis, University of Göttingen, 1987.
- [116] Burkhard Will. Schrittweisensteuerung bei Verfahren zur Lösung von Minimax-Aufgaben. Diploma thesis, University of Göttingen, 1982.

Name Florian Boßmann
Address Institute for Numerical and Applied Mathematics
University of Göttingen
Lotzestraße 16-18
37083 Göttingen / Germany
Email: f.bossmann@math.uni-goettingen.de

Personal Details

Gender Male
Date of birth 9th of June, 1986
Place of birth Emmerich, Germany
Citizenship German

Education

07/2010-10/2013 Ph.D. student of mathematics at the University of Göttingen (Germany). Supervisor: G. Plonka-Hoch.
01/2010-06/2010 Ph.D. student of mathematics at the University of Duisburg (Germany). Supervisor: G. Plonka-Hoch.
10/2005-12/2009 Student of mathematics (for diploma) at the University of Duisburg (Germany) with minor subject computer science. Sept. 20, 2007 Intermediate examination for diploma in mathematics. Dec. 18, 2009 Examination (Diplom) in mathematics (passed with distinction). Thesis title: “Entwicklung einer automatisierten Auswertung von bildgebenden Ultraschallverfahren”, Advisor: G. Plonka-Hoch.
08/2002-06/2005 Vocational collage “Berufskolleg Wesel” in Wesel (Germany). Jun. 2005 Fachhochschulreife, assistant for Business Informatics.
08/1996-07/2002 Secondary school “Gymnasium Aspel” in Rees (Germany).

Research Experience

Since 09/2013 Assistant coordinator of the joint research project “Efficient analysis of high dimensional ultrasonic data in non-destructive material testing (ZeMat)” supported by the BMBF, research fellow of subproject 1 “Reconstruction of 3D defect characteristics in welds”, Göttingen.
11/2011-08/2013 Research fellow of SFB 755 (Nanoscale Photonic Imaging), subproject C11 “Fresnel wavelets for coherent diffractive imaging”, Göttingen.
01/2010-10/2011 Research fellow of the project “Model based image pattern recognition in non-destructive material testing” supported by the BMWI in collaboration with Salzgitter Mannesmann Forschung GmbH, Duisburg/Göttingen.
